

Nuevos métodos proteómicos basados en ultrasonidos y
espectrometría de masas para la búsqueda de biomarcadores en
enfermedades reumáticas

Tesis Doctoral

Carolina Fernández Costa

Octubre 2013

Tesis doctoral dirigida por el Dr. Francisco Javier Blanco García, el Dr. José Luis
Capelo Martínez y la Dra. Cristina Ruiz Romero.

Grupo de Proteómica – Servicio de Reumatología- Instituto de Investigación
Biomédica de A Coruña.

Los directores de esta tesis doctoral, D. Francisco Javier Blanco García, D. José Luis Capelo Martínez y Dña. Cristina Ruiz Romero,

CERTIFICAN:

Que Dña. Carolina Fernández Costa, Licenciada en Ciencias Químicas por la Universidad de Vigo, ha realizado en el Servicio de Proteómica del Instituto de Investigación Biomédica de A Coruña (INIBIC), y bajo su dirección, el trabajo “Nuevos métodos proteómicos basados en ultrasonidos y espectrometría de masas para la búsqueda de biomarcadores en enfermedades reumáticas”, el cual, reúne todas las condiciones para ser presentado como Tesis Doctoral.

Y para que así conste, firman el presente certificado en A Coruña, a 23 de Septiembre de 2013.

Dr. Francisco J. Blanco García
Director

Dr. José Luis Capelo Martínez
Director

Dra. Cristina Ruiz Romero
Directora

D. Isaac Fuentes Boquete, tutor de esta tesis,

CERTIFICA:

Que Dña. Carolina Fernández Costa, Licenciada en Ciencias Químicas por la Universidad de Vigo, ha realizado en el Servicio de Proteómica del Instituto de Investigación Biomédica de A Coruña (INIBIC), y bajo su tutela, el trabajo “Nuevos métodos proteómicos basados en ultrasonidos y espectrometría de masas para la búsqueda de biomarcadores en enfermedades reumáticas”, el cual reúne todas las condiciones para ser presentado como Tesis Doctoral.

Y para que así conste, firmo el presente certificado en A Coruña, a 23 de Septiembre de 2013.

Dr. Isaac Fuentes Boquete
Tutor

Agradecimientos

Hay mucha gente que de alguna manera ha participado en el desarrollo de esta tesis, a todos ellos les transmito mi más sincero agradecimiento.

Me gustaría comenzar agradeciendo en primer lugar a mis directores de tesis, el Dr. Francisco J. Blanco, el Dr. José Luis Capelo, y la Dra. Cristina Ruiz, por haber confiado en mí para realizar todo este trabajo, y haberme sabido guiar con su experiencia durante todos estos años.

A mis compañeros del grupo de proteómica, Valentina, Patricia, Jesús, Bea y Lucía, tengo que agradecerles todo la ayuda que me han brindado y lo bien que me han soportado durante estos años, y sé que no es fácil. En especial, le agradezco a Valentina el estar siempre dispuesta a ayudarnos y enseñarnos todo lo que sabe dedicándonos el tiempo que sea necesario. Gracias por vivir con tanta pasión la ciencia. A Lu y a Bea, mis compañeras y amigas, por sufrir conmigo todas las batallas del día a día dentro y fuera del trabajo.

A todos los compañeros del INIBIC, por haberme ayudado siempre en todo lo que he necesitado. En especial al grupo de Genómica, a Nacho y a Sara, por su paciencia con las “muestras de suero”, y a Juan por todo el tiempo que invirtió en mis “picos de masas”.

A Miguel y a Florentino, por todo el trabajo que realizaron con los datos de los espectros de masas.

Por supuesto, a mis padres y a mis dos hermanos, os agradezco el haber estado a mi lado en los momentos difíciles y haberme animado siempre a seguir adelante. A mi familia en general y a Lorena, por el apoyo y cariño que me dais.

A Ángel, por aguantarme día a día, y hacerme sonreír hasta en los peores momentos. Gracias por creer en mí.

Índice

Índice de abreviaturas	13
Índice de figuras	15
Índice de tablas	16
Resumen	17
1. Introducción	25
1.1. Enfermedades reumáticas	27
1.2. Artrosis	29
1.2.1. Prevalencia e impacto económico	29
1.2.2. Factores de riesgo	31
1.2.3. Etiología de la artrosis	34
1.2.4. El diagnóstico y sus limitaciones	38
1.3. Proteómica	40
1.3.1. Búsqueda de biomarcadores	43
1.3.2. Estrategias proteómicas en la búsqueda de biomarcadores	47
1.3.3. Métodos de depleción	48
1.3.4. Métodos de digestión de proteínas	49
1.3.5. Métodos de separación basados en gel	51
1.3.6. Métodos de separación cromatográficos	53
1.3.7. Espectrometría de masas	57
1.3.7.1. Fuentes de iones	58
1.3.7.2. Analizadores de masas	61
1.3.7.3. Identificación de proteínas	65
2. Objetivos	67
3. Capítulos	71
3.1. Artículo I: Comparación de la depleción y la ecualización para reducir las proteínas más abundantes en suero humano	73
3.2. Artículo II: Depleción secuencial de suero humano en la búsqueda de biomarcadores	85
3.3. Artículo III: Evaluación de los ultrasonidos indirectos de alta intensidad en la digestión de proteomas complejos con tripsina inmovilizada	99
3.4. Capítulo IV: Depleción secuencial acoplada a extracción secuencial en C18 como herramienta rápida para la búsqueda de biomarcadores de enfermedades reumáticas en suero humano	109
3.4.1. Material y métodos	113
3.4.2. Resultados	119

4. Discusión	129
4.1. Discusión artículo I	131
4.2. Discusión artículo II	137
4.3. Discusión artículo III	141
4.4. Discusión capítulo IV	145
4.5. Consideraciones finales	149
5. Conclusiones	153
6. Bibliografía	157
7. Anexos	169
7.1. Anexo I: Material suplementario del artículo I	171
7.2. Anexo II: Material suplementario del artículo II	181
7.3. Anexo III: Material suplementario del artículo III	187
7.4. Anexo IV: Material suplementario del capítulo IV	191
7.5. Anexo V: Artículo IV. Identification of a Panel of Novel Serum Osteoarthritis Biomarkers	193

Índice de abreviaturas

2-DE: electroforesis bidimensional (2-dimensional electrophoresis).
2D-DIGE: electroforesis diferencial en gel (2-Dimensional Difference Gel Electrophoresis).
2D-GE: electroforesis bidimensional en gel (2 dimensional-gel electrophoresis).
ACN: acetonitrilo.
ACR: Colegio Americano de Reumatología (American College of Rheumatology).
ADN: ácido desoxirribonucleico.
ADN-mt: ácido desoxirribonucleico mitocondrial.
AINEs: antiinflamatorios no esteroideos.
AR: artritis reumatoide.
ARN: ácido ribonucleico.
BCA: ácido bicinconínico.
BMP: proteínas morfogenéticas óseas (Bone Morphogenetic Proteins).
CHCA: ácido α -ciano-4-hidroxicinámico.
CID: disociación inducida por colisión (Collision-Induced Dissociation).
DHB: ácido 2,5-dihidroxibenzoico.
DTT: ditiotreitól.
ESI: ionización por electronebulización o electrospray (ElectroSpray Ionization).
FM: fibromialgia.
FT: transformada de Fourier.
GDF5: factor 5 de crecimiento y diferenciación.
GWAS: estudios de asociación del genoma completo (Genome Wide Association Study).
HILIC: cromatografía de interacciones hidrofílicas (Hydrophilic Interaction Chromatography).
HPLC: cromatografía de líquidos de alta eficacia (High Efficiency Liquid Chromatography).
HPT: haptoglobina.
HUPO: Organización del Proteoma Humano (Human Proteome Organization).
ICR: resonancia iónica ciclotrónica (Ion Cyclotron Resonance).
IEX: intercambio iónico (Ion Exchange).
IHI-UE: energía de ultrasonidos de alta intensidad indirecta (Indirect High Intensity-Ultrasonic Energy).

IL: interleucina.

IPG: gradiente de pH inmovilizado (Immobilized pH Gradient).

IR: infrarrojos.

LC: cromatografía líquida (liquid chromatography).

m/z: relación masa/carga.

MALDI: desorción/ionización por láser asistida por matriz (Matrix-Assisted Laser Desorption Ionization).

mg: miligramos.

ml: mililitros.

MMP: metaloproteasa de matriz (Matrix MetalloProteinase).

MS: espectrometría de masas (Mass Spectrometry).

NFG: factor de crecimiento nervioso (Nerve Growth Factor).

nm: nanómetros.

OA: artrosis (Osteoarthritis).

PAGE: electroforesis en geles de poliacrilamida (polyacrylamide gel electroforesis).

PCA: análisis de componentes principales (Principal Component Analysis).

pg: picogramos.

pI: punto isoeléctrico.

PTM: modificaciones postraduccionales (post-translational modifications).

RP: fase reversa (reverse phase).

SDS: dodecil sulfato sódico (sodium dodecyl sulfate).

SER: Sociedad Española de Reumatología.

TGF: factor de crecimiento transformador (Transforming Growth Factor).

TIMP: inhibidores de proteasas (Tissue Inhibitor of Metalloproteinase).

TNF: factor de necrosis tumoral (Tumor Necrosis Factor).

TOF: tiempo de vuelo (Time Of Flight).

US: ultrasonidos.

UV-VIS: ultravioleta visible.

µg: microgramo.

Índice de figuras

Figura 1. Articulaciones sinoviales del cuerpo humano.	27
Figura 2. Imágenes representativas de una articulación sana y artrósica.	29
Figura 3. Gráfica en la que se representan los principales factores de riesgo asociados con la artrosis.	31
Figura 4. Estructura del cartílago articular hialino, donde se muestran las diferentes capas en las que se divide, así como el único componente celular del cartílago, el condrocito.	35
Figura 5. Moléculas que intervienen en el equilibrio entre los procesos anabólicos y catabólicos del cartílago.	36
Figura 6. Imagen que representa el paso del genoma al proteoma.	40
Figura 7. Imagen que representa la complejidad del proteoma en comparación con el genoma, y los factores que influyen en síntesis de las proteínas a partir de los genes.	41
Figura 8. Principales aplicaciones de la Proteómica.	43
Figura 9. Rango dinámico de concentración de las proteínas del plasma	45
Figura 10. Distribución de las proteínas más abundantes del plasma.	46
Figura 11. Estrategias proteómicas empleadas en la búsqueda de biomarcadores en muestras de suero.	48
Figura 12. Método de depleción basado en inmunoafinidad.	49
Figura 13. Gel del 10% de acrilamida teñido con plata, en el que se separaron las proteínas de diferentes muestras de suero.	51
Figura 14. Proceso de una separación de proteínas mediante la técnica de electroforesis bidimensional.	52
Figura 15. Esquema representativo de la técnica de electroforesis diferencial en gel (2D-DIGE) para la separación y cuantificación de proteínas.	53
Figura 16. Tipos de cromatografía de líquidos más empleados en la separación de proteínas.	54
Figura 17. Componentes de un espectrómetro de masas.	58
Figura 18. Esquema de la ionización por electronebulización.	59
Figura 19. Desorción/ionización por láser asistida por matriz (MALDI).	60
Figura 20. Tipos de analizadores de masas más empleados en proteómica.	61

Figura 21. Espectrómetro de masas cuadrupolar.	62
Figura 22. Espectrómetro de masas de trampa de iones.	63
Figura 23. Espectro de masas de la proteína albúmina digerida con tripsina y analizada en un espectrómetro de masas con desorción/ionización con láser asistida por matriz y analizador de tiempo de vuelo (MALDI-TOF).	65

Índice de tablas

Tabla 1. Distribución de la prevalencia de las enfermedades reumáticas seleccionadas en la población española de más de 20 años de edad, estratificada por sexo.	30
Tabla 2. Visitas médicas realizadas en España debido a enfermedades musculoesqueléticas, clasificadas por el tipo de enfermedad, por consumo de AINEs y por las compensaciones económicas recibidas por discapacidad.	31
Tabla 3. Criterios clínicos del Colegio Americano de Reumatología para la clasificación de la artrosis.	38
Tabla 4. Biomarcadores de artrosis empleados en ensayos clínicos.	39
Tabla 5. Proteasas empleadas en la digestión de proteínas.	50
Tabla 6. Principales matrices para MALDI empleadas en proteómica.	60
Tabla 7. Características de los espectrómetros de masas híbridos más empleados en proteómica.	64

RESUMEN

Resumen

La artrosis es la enfermedad articular más frecuente, está caracterizada por la degradación del cartílago articular y crea dolor, rigidez e incapacidad funcional. A pesar de su elevada prevalencia, los métodos diagnósticos actualmente disponibles son limitados y poco sensibles. Se basan en la descripción subjetiva de los síntomas del paciente, en la rigidez articular y en las pruebas radiológicas. Estas limitaciones conllevan, generalmente, a un primer diagnóstico cuando ya se han producido daños en el cartílago. Para evitar llegar a este daño irreversible de la articulación, la investigación en este campo se está centrando en la búsqueda de marcadores biológicos que faciliten el diagnóstico temprano de la enfermedad.

El objetivo de esta tesis fue desarrollar tecnologías proteómicas para la búsqueda de biomarcadores proteicos que mejoren el diagnóstico, pronóstico y tratamiento de la artrosis. Nuestra investigación se focalizó en el desarrollo de metodologías rápidas y económicas para buscar biomarcadores en suero mediante espectrometría de masas.

Inicialmente se realizaron diversos experimentos para encontrar un nuevo método de depleción de proteínas abundantes del suero. Para ello se comenzó con un estudio comparativo de varias metodologías ya existentes para deplecionar proteínas mediante precipitaciones químicas y kits comerciales. A continuación se desarrolló un nuevo método de depleción basado en el estudio anterior, consistente en dos precipitaciones químicas secuenciales. Para concluir este estudio, se comprobó la utilidad del método de depleción secuencial en la búsqueda de biomarcadores de artrosis en sueros de pacientes artrósicos de grado IV y en sueros de donantes control, mediante la técnica de electroforesis bidimensional diferencial en gel (2D-DIGE). Se identificaron 16 proteínas moduladas en artrosis con respecto a los controles mediante espectrometría de masas, y se realizó la validación de una de ellas mediante inmunodetección.

Con la finalidad de simplificar los espectros de masas y mejorar las identificaciones de las proteínas, se realizó otro estudio para aplicar los ultrasonidos indirectos de alta eficacia con tripsina inmovilizada en la digestión de las proteínas. Los resultados obtenidos demostraron que los ultrasonidos indirectos no favorecen la

homogenización del enzima en la muestra, con lo que la digestión resultó menos eficaz que con el enzima en solución.

Por último, en base a los anteriores estudios, se desarrolló un nuevo método rápido de búsqueda de marcadores peptídicos en suero. El método consistió en un primer paso de depleción secuencial, seguido de la digestión proteica con tripsina en solución acelerada por ultrasonidos indirectos. A continuación se realizó una separación de los péptidos mediante extracción en fase sólida con octadecilo (C18) inmovilizado, y por último se analizaron los péptidos por espectrometría de masas. Esta metodología se aplicó a muestras de pacientes con enfermedades reumáticas (artrosis, artritis reumatoide y artritis psoriásica) y donantes control. Los resultados obtenidos proporcionaron un perfil peptídico capaz de clasificar los sueros en las cuatro condiciones estudiadas.

Resumo

A artrose é a enfermidade articular máis frecuente, está caracterizada pola degradación da cartilaxe articular creando dor, rixidez e incapacidade funcional. Malia a súa elevada prevalencia, os métodos diagnósticos actualmente dispoñibles son limitados e pouco sensibles. Baséanse na descrición subxectiva dos síntomas do paciente, na rixidez articular e nas probas radiolóxicas. Estas limitacións implican, xeralmente, un primeiro diagnóstico cando xa se produciron danos na cartilaxe. Para evitar chegar a este dano irreversible da articulación, a investigación neste campo estase focalizando na procura de marcadores biolóxicos que faciliten o diagnóstico prematuro da enfermidade.

O obxectivo desta tese foi desenvolver tecnoloxías proteómicas para a procura de biomarcadores proteicos que melloren o diagnóstico, prognóstico e tratamento da artrose. A nosa investigación focalizouse no desenvolvemento de metodoloxías rápidas e económicas para buscar biomarcadores en soro mediante espectrometría de masas.

Inicialmente realizáronse diversos experimentos para atopar un novo método de depleción de proteínas abundantes do soro. Para iso comezouse cun estudo comparativo de varias metodoloxías xa existentes para deplecionar proteínas mediante precipitacións químicas e kits comerciais. A continuación desenvolveuse un novo método de depleción baseado no estudo anterior, consistente en dúas precipitacións químicas secuenciais. Para concluír este estudo, comprobouse a utilidade do método de depleción secuencial na procura de biomarcadores de artrose en soros de pacientes con artrose de grado IV e en soros de doantes control, mediante a técnica de electroforese bidimensional diferencial en xel (2D-DIGE). Identificáronse 16 proteínas moduladas en artrose con respecto aos controis mediante espectrometría de masas, e realizouse a validación dunha delas mediante inmunodetección.

Coa finalidade de simplificar os espectros de masas e mellorar as identificacións das proteínas, realizouse outro estudo para aplicar os ultrasóns indirectos de alta eficacia con tripsina inmovilizada na dixestión das proteínas. Os resultados obtidos demostraron que os ultrasóns indirectos non favorecen a homoxeneización do

enzima na mostra, co que a dixestión resultou menos eficaz que co enzima en solución.

Para rematar, en base aos anteriores estudos, desenvolveuse un novo método rápido de procura de marcadores peptídicos en soro. O método consistiu nun primeiro paso de depleción secuencial, seguido da dixestión proteica con tripsina en solución acelerada por ultrasóns indirectos, a continuación realizouse unha separación dos péptidos mediante extracción en fase sólida con octadecilo (C18) inmovilizado e para rematar analizáronse os péptidos por espectrometría de masas. Esta metodoloxía aplicouse a mostras de pacientes con enfermidades reumáticas (artrose, artrite reumatoide e artrite psoriásica) e de doantes control. Os resultados obtidos proporcionaron un perfil peptídico capaz de clasificar os soros nas catro condicións estudadas.

Abstract

Osteoarthritis (OA) is the most common articular disease characterized by cartilage degradation, pain, stiffness and loss of function. Despite its high prevalence, the diagnostic methods currently available are limited and lack sensitivity. Nowadays, the diagnosis of OA relies on the subjective description of pain symptoms, stiffness in the affected joints, and radiography. For this reason, in most cases an extensive deterioration of cartilage already exists at the time of diagnosis. To avoid the limitations of conventional tests, there is an increasing interest in identifying new specific biological markers for early OA diagnosis.

The aim of this work was the development of new proteomics strategies for the discovery of protein biomarkers to improve OA diagnosis, prognosis and treatment. Our research has been focused on the generation of fast and cheap methodologies to find biomarkers of rheumatic diseases in serum samples by mass spectrometry (MS) analysis.

Firstly, several experiments were performed to seek a new depletion method for abundant proteins in serum samples. A comparative study of existing depletion methodologies (both chemicals and commercial kits) was performed in order to develop a new depletion method based on these previous results. To conclude this study, we verified the utility of the sequential depletion method in the search of OA biomarkers in serum samples obtained from IV grade OA patients and controls, using the two dimensional difference gel electrophoresis (2D-DIGE) technique. 16 differentially expressed proteins were identified in OA samples compared to controls, and one of them was validated by immunodetection.

The indirect high intensity ultrasonic energy (IHI-UE) using immobilized trypsin was applied in order to simplify the mass spectra and improve the protein identification. The results showed that protein digestion with the immobilized trypsin was less effective compared to in-solution digestion after application of IHI-UE.

Finally, and according to the previous studies, a new fast method to search peptide markers in serum samples was developed. This method consisted of a sequential depletion step followed by in-solution protein digestion with trypsin using IHI-UE. Then, the resulting peptides were separated by solid phase extraction and analyzed by MS. Samples from controls and patients affected by rheumatic pathologies (osteoarthritis, rheumatoid arthritis and psoriatic arthritis) were analyzed by this methodology. The results obtained provided different peptidic profiles of the pathologies, which allowed us to classify them according to these profiles.

1. INTRODUCCIÓN

1. Introducción:

1.1. Enfermedades reumáticas

El término “enfermedades reumáticas” engloba un conjunto de enfermedades muy diversas y complejas, de origen no traumático, que afectan principalmente al aparato locomotor o musculoesquelético (articulaciones, huesos, músculos, tendones y ligamentos), pero que también pueden afectar a otros sistemas u órganos como el corazón, los pulmones, los ojos, la piel, los vasos sanguíneos y otros tejidos conectivos que se encuentran en todo el organismo, de ahí su diversidad y complejidad.

La palabra “reuma” proviene del griego y significa “flujo”. Antiguamente, en la época en que la medicina se basaba en la “teoría de los humores”, se creía que las enfermedades reumáticas aparecían por el “flujo (o reuma) de un humor” desde el cerebro hacia las articulaciones. Era la explicación que la medicina de aquella época daba al dolor y a la inflamación que aparecen en las enfermedades reumáticas.

Con el tiempo, y al igual que la medicina, este concepto evolucionó, y hoy día el término “reuma”, aunque se use coloquialmente de forma inespecífica para referirnos al dolor y/o a la inflamación de las articulaciones, en realidad carece de significado ya que el “reuma” como tal, no existe. Hoy en día existen las enfermedades reumáticas o reumatismos.

Existen más de 100 enfermedades reumáticas distintas, como la artrosis, la gota, la artritis reumatoide, la osteoporosis o la fibromialgia entre

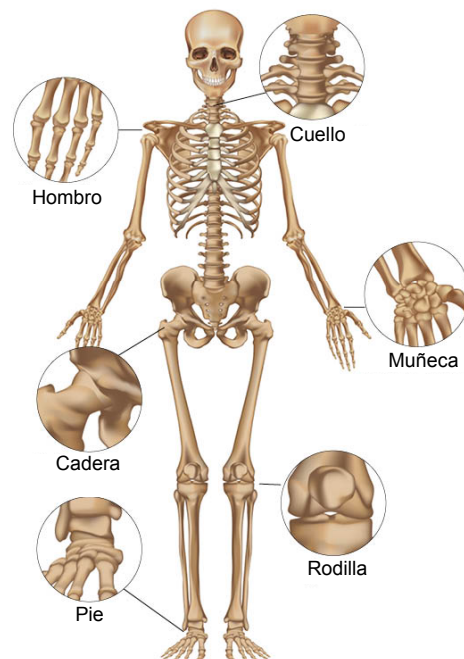


Figura 1. Articulaciones sinoviales del cuerpo humano. Figura adaptada de la página <http://healthpages.org>

otras, que tienen diferente origen. Por ejemplo, la artritis reumatoide se caracteriza por la inflamación de las articulaciones debida en parte a procesos autoinmunes. La gota es una artritis debida al depósito de cristales microscópicos sobre las articulaciones. En cambio, la artrosis conlleva desgaste y degeneración articular sin necesidad de que se produzca inflamación. Por su parte, la osteoporosis se debe a la pérdida de masa ósea y a la alteración de la disposición de las trabéculas óseas. A pesar de todas las diferencias que existen entre las enfermedades reumáticas, todas ellas presentan en general síntomas de dolor, rigidez articular y pérdida de capacidad funcional. En la figura 1 están representadas las articulaciones sinoviales que se pueden ver afectadas en estas enfermedades.

Las enfermedades reumáticas se encuentran entre las dolencias que afectan con mayor frecuencia a los seres humanos. Según el estudio que realizó la Sociedad Española de Reumatología en el año 2000 (EPISER), el 18 % de los mayores de 20 años padecen algún trastorno reumático. Dentro de las enfermedades reumáticas, la artrosis es la patología más frecuente en España.

El tratamiento de las enfermedades reumáticas ha mejorado mucho en los últimos años. La gota se puede tratar hoy en día con medicamentos que impiden la producción de ácido úrico por nuestro organismo. Las artropatías inflamatorias (artritis reumatoide, artritis psoriásica, espondilitis anquilopoyética) pueden ser controladas y conducidas hacia la remisión de su actividad clínica mediante el empleo adecuado de fármacos como el metotrexato o la leflunomida, o con las terapias biológicas desarrolladas más recientemente. Para la osteoporosis se dispone de medicamentos capaces de reducir las fracturas producidas por la enfermedad. Por el contrario, para la artrosis no hay tratamientos efectivos y, a pesar de toda la investigación realizada hasta el momento, a día de hoy el tratamiento de la misma sólo es sintomático. Por ello, en esta tesis nos centramos principalmente en el estudio de esta enfermedad con el fin de facilitar su diagnóstico y seguimiento preciso, que son requisitos imprescindibles para el desarrollo de terapias alternativas que puedan posibilitar la curación.

1.2. Artrosis.

La artrosis es una patología articular degenerativa caracterizada por alteración en la integridad del cartílago y el hueso subcondral, así como alteraciones morfológicas y biomecánicas de la membrana sinovial y la cápsula articular [1, 2]. En la figura 2 se pueden observar las diferencias entre una articulación sana y una articulación artrósica.

Actualmente se ha pasado a considerar a la artrosis no como un única enfermedad, sino como un grupo heterogéneo de patologías con manifestaciones clínicas similares y cambios patológicos y radiológicos comunes [3]. De acuerdo

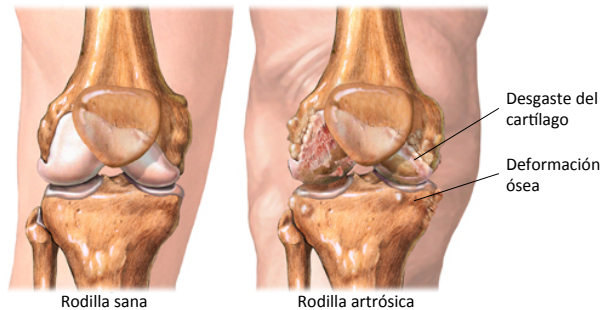


Figura 2. Imágenes representativas de una articulación sana y artrósica. Figura adaptada de A.D.A.M. Images.

con esta línea, el Subcomité de Artrosis del Colegio Americano de Reumatología (American College of Rheumatology, ACR) definió la artrosis como un grupo heterogéneo de condiciones que conducen a síntomas y signos articulares, los cuales se asocian con defectos en la integridad del cartílago articular, además de cambios relacionados con el hueso subcondral y con los márgenes articulares [4]. La artrosis se puede clasificar según su etiología en primaria (o idiopática) o secundaria a otras enfermedades como las metabólicas, endocrinas, óseas y traumatismos entre otras.

1.2.1. Prevalencia de la artrosis e impacto económico

La artrosis es la artropatía más frecuente en la población, afecta fundamentalmente a los mayores de 50 años. En sociedades industrializadas, la artrosis es la principal causa de inhabilitación física. Se han realizado multitud de estudios epidemiológicos de la artrosis, sin embargo, no es sencillo determinar la prevalencia de la enfermedad. El principal problema se encuentra en la falta de consenso entre los diferentes autores a la hora de definir el caso de artrosis y las metodologías empleadas.

En el año 2001, la Sociedad Española de Reumatología (SER) realizó un estudio de la prevalencia de las enfermedades reumáticas en España llamado EPISER. En la tabla 1 se muestran los resultados de dicho estudio y se puede comprobar la elevada prevalencia de la artrosis frente a las otras enfermedades reumáticas, así como la mayor incidencia en el género femenino. Se calcula que al menos una de cada cinco personas mayores de 16 años padece alguna dolencia reumática. Este dato se incrementa con la edad, llegando a afectar a una de cada dos a partir de los 65 años [5, 6]. Aproximadamente un 16 % de la población mayor de 20 años sufre artrosis, de acuerdo con los informes de prevalencia que realiza periódicamente la Sociedad Española de Reumatología (SER).

Enfermedad	% Mujeres	% Hombres	% Prevalencia Total
Artrosis de rodilla	14	5.7	10.2
Artrosis de manos	9.5	2.3	6.2
Artritis reumatoide	0.8	0.2	0.5
Dolor lumbar	17.8	11.3	14.8
Fibromialgia	4.2	0.2	2.4

Tabla 1. Distribución de la prevalencia de las enfermedades reumáticas seleccionadas en la población española de más de 20 años de edad estratificada por sexo. Datos pertenecientes al estudio EPISER, publicado por Carmona L. y colaboradores [5].

Esta patología conlleva una carga económica elevada para la familia y la sociedad. Su coste global se estima en un 2 % del producto interior bruto (PIB) para los países de nuestro entorno [7-9]. En el estudio EPISER calculan que un 33 % de la población había visitado al médico en un año por problemas musculoesqueléticos. En la tabla 2 podemos observar el elevado número de visitas realizadas por pacientes con artrosis de rodilla, un 66 %, y el porcentaje de población que consume antiinflamatorios no esteroideos (AINEs). Todo ello implica un elevado coste económico que podemos ver reflejado en el estudio ArtRoCad para la evaluación de la utilización de los recursos sanitarios y de la repercusión socioeconómica de la artrosis de rodilla y cadera, realizado por la SER, en donde se calcula un coste medio anual por paciente de 1502 € y un coste total de casi 5000 millones de euros [10].

Enfermedad	Consultaron al médico (%)	Consumieron AINEs (%)	Recibieron compensaciones económicas (%)
OA rodilla	66.4	45.7	5.4
OA manos	58.8	38.2	2.2
AR	72.7	63.6	9.1
Dolor lumbar	61.2	40.9	8.0
FM	76.9	55.8	7.7

Tabla 2. Visitas médicas realizadas en España debido a enfermedades musculoesqueléticas, clasificadas por el tipo de enfermedad, por consumo de AINEs y por las compensaciones económicas recibidas por discapacidad. OA: artrosis; AR: artritis reumatoide; FM: fibromialgia; AINEs: anti-inflamatorios no esteroideos. Datos pertenecientes al estudio EPISER, publicado por Carmona L. y colaboradores [5].

1.2.2. Factores de riesgo

La artrosis es una enfermedad de origen multifactorial e influyen en ella tanto factores de riesgo endógenos como exógenos. Dentro de los factores endógenos se encuentran la edad, el sexo, la raza, la genética y la densidad ósea. Entre los factores exógenos se encuentran la obesidad, los traumatismos, la cirugía articular, el deporte, y el tipo de profesión [11]. En la figura 3 están representados los factores de riesgo más relevantes en esta patología.

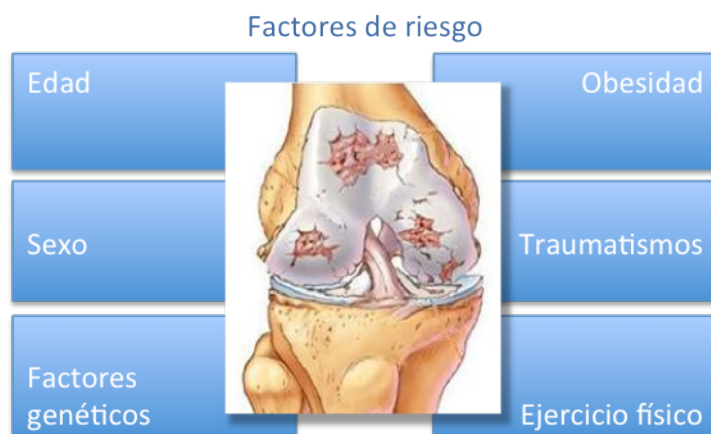


Figura 3. Gráfica en donde están representados los principales factores de riesgo asociados con la artrosis.

Edad

La edad es el factor de riesgo más fuertemente relacionado con la artrosis [12]. Más del 80 % de los mayores de 75 años sufren artrosis. Tanto la prevalencia como la incidencia de la enfermedad se incrementan con el aumento de la edad en cualquier articulación que se considere. Este factor parece estar relacionado con la existencia de cambios anatómicos, estructurales y bioquímicos que favorecerían la degradación del cartílago articular en las personas mayores. A pesar de ser una enfermedad relacionada con la edad, la artrosis no es una consecuencia inevitable del envejecimiento.

Sexo

Existe un claro aumento en la prevalencia de la artrosis en mujeres mayores de 50 años con respecto al género masculino. Estas diferencias entre géneros a partir de los 50 años parecen ser debidas a la deficiencia de estrógeno en las mujeres postmenopáusicas. Los condrocitos articulares poseen receptores de estrógeno, de manera que dicha hormona podría regular el funcionamiento estas células. En estudios recientes se ha relacionado la terapia de sustitución de estrógenos con la disminución en el riesgo de padecer artrosis de rodilla y cadera en mujeres postmenopáusicas [13, 14].

Factores genéticos

Diversos estudios han demostrado la importancia de los factores hereditarios en la artrosis. Los resultados obtenidos en estudios de heredabilidad en gemelos, parejas de hermanos y familias, estiman que los factores genéticos representan un 50% del riesgo de desarrollar artrosis de cadera o rodilla [15, 16]. Hasta no hace mucho, sólo se habían encontrado tres posiciones fijas del cromosoma (o *loci*) (*GDF5*, cromosoma 7q22, and *MCF2L*) significativamente asociados a la artrosis. El factor de crecimiento y diferenciación 5 (*GDF5*) es un ligando de la superfamilia TGF-beta (factor de crecimiento transformador) que está muy relacionado con las proteínas morfogenéticas de hueso (BMP). Tiene un papel crítico en la formación de las articulaciones sinoviales y también se encuentra sobre expresado en la reparación de las fracturas. Una disminución en su expresión limita la función

reparadora del cartílago dañado [17]. En un estudio reciente del año 2011 confirmaron la relación de 6 genes del cromosoma 7q22 con el riesgo de padecer artrosis de rodilla (*PRKAR2B*, *HPB1*, *COG5*, *GPR22*, *DUS4L* y *BCAP29*) [18]. Aunque estos genes ya se habían relacionado anteriormente con la artrosis, aún no se han podido explicar las rutas implicadas en esta asociación. También se relacionan con el riesgo de padecer artrosis algunos polimorfismos en el gen *MCF2L*. Este gen regula la neurotrofina 3, que pertenece a la familia de los factores de crecimiento nerviosos (NFG). El tratamiento en pacientes con OA de rodilla con un anticuerpo que inhibe NFG se relaciona con una reducción del dolor articular y una mejora del funcionamiento articular [19].

En 2012 se publicó el resultado final del estudio completo arcOGEN GWAS (estudios de asociación del genoma completo) realizado por el Consorcio de Reino Unido de Investigación Genética en OA. Encontraron 8 nuevos *loci* asociados con el riesgo de desarrollar artrosis. La asociación más fuerte se encontró en el cromosoma 3, en concreto en el gen *GNL3* que codifica para la nucleostemina, que se encuentra elevada en los condrocitos OA respecto a los controles. En el cromosoma 9 encontraron el gen que codifica para *ASTN2* (astrotactina 2). En el cromosoma 6 encontraron *FILIP1* (proteína 1 de interacción con filamina A) y *SENP6* (peptidasa específica de sentrina). También encontraron que el gen *CHST11* (carbohidrato sulfotransferasa 11) se encuentra elevado en cartílago OA frente al control, dicho gen tiene un papel relevante en el desarrollo esquelético, rutas de señalización y progresión del cáncer. Otro gen fuertemente relacionado con artrosis es el *FTO*, que está relacionado con la materia grasa y la obesidad, de manera que actúa en la regulación del peso corporal [20].

Con todos los resultados obtenidos hasta ahora, hay un total de 11 *loci* relacionados con el riesgo de desarrollar artrosis en la población europea. Se han encontrado más genes en diversos estudios, pero muchos de ellos no se han confirmado en trabajos posteriores.

En otros estudios genéticos se han encontrado unos polimorfismos heredados por vía materna en el ADN mitocondrial (ADN-mt), denominados haplogrupos de ADN-mt, que influyen en la eficacia energética de la mitocondria. Se ha demostrado recientemente que el haplogrupo mitocondrial J es un factor protector frente al desarrollo de la artrosis de rodilla y cadera, dando lugar a una menor destrucción del cartílago articular en una subpoblación española [21-25].

Obesidad

Estudios recientes demuestran que la obesidad precede al desarrollo de la artrosis [26, 27]. Las personas que tienen sobrepeso tienen una alta prevalencia de artrosis de rodilla. Además, en las personas con artrosis, tener sobrepeso aumenta el riesgo de progresión radiográfica de la enfermedad [28, 29]. En la mayoría de los estudios, el aumento del riesgo de artrosis de rodilla por sobrepeso es mayor en mujeres que en hombres [11].

Traumatismos

Los traumatismos sobre las articulaciones pueden conducir a la aparición posterior de una artrosis. En el estudio de Framingham [30] se demostró que un traumatismo previo sobre la rodilla tiene un riesgo 5-6 veces superior de artrosis de rodilla en el hombre y 3 veces en la mujer. En otro estudio también se encontró un alto riesgo de artrosis de rodilla en sujetos que habían sufrido un traumatismo articular en su juventud [31].

Ejercicio físico

Se ha encontrado un aumento en el riesgo de padecer artrosis en personas que practican deportes de alta intensidad, con impacto directo en las articulaciones. También en deportes de equipo en donde se puede producir contacto entre jugadores, como los jugadores de fútbol [32]. Considerando el ejercicio físico que deriva de la realización de una actividad profesional, hay determinados trabajos que implican tareas repetitivas que sobrecargan las articulaciones y producen fatiga en los músculos que protegen las articulaciones [33].

1.2.3. Etiología de la artrosis

La artrosis actualmente se considera una enfermedad de la articulación en su conjunto. El fenómeno que caracteriza la artrosis es la destrucción del cartílago articular junto con esclerosis del hueso subcondral, formación de osteofitos en los márgenes del hueso e inflamación crónica no específica de la membrana sinovial. La capacidad del cartílago para soportar las presiones generadas durante la carga

reside en las complejas interacciones entre los distintos componentes de su matriz extracelular, sintetizadas por un número muy pequeño de células, los condrocitos.

Composición del cartílago

El cartílago articular está formado por un componente celular y una matriz extracelular. Las células presentes en el cartílago se denominan condrocitos, representan el 1-2 % del volumen total del cartílago. La matriz extracelular está compuesta por un 65-80 % de agua, un 10-30 % de colágeno, un 5-10 % de proteoglicanos y un 15-20 % de otras proteínas de matriz. En la figura 4 se puede ver la disposición de los componentes del cartílago.

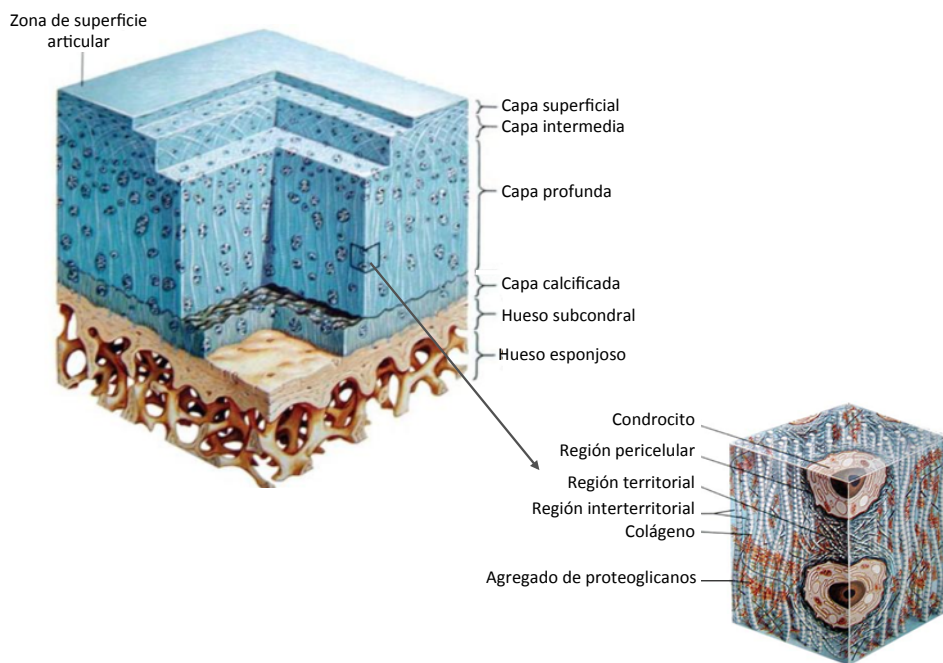


Figura 4. Estructura del cartílago articular hialino, mostrando las diferentes capas en las que se divide, así como el único componente celular del cartílago, el condrocito. Imagen adaptada de ELSEVIER.INC.- NETTERIMAGES.COM .

El colágeno del cartílago es del tipo II en un 90-95 %. Se dispone en forma de una red tridimensional de fibras dando la forma, volumen y la fuerza tensora del cartílago. El otro 10 % corresponde a colágenos tipos VI, IX, XI, XII y XIV [34]. Los proteoglicanos son macromoléculas altamente hidrofílicas y cargadas negativamente y se encuentran absorbidos en el interior de la red de colágeno

creando presión al retener gran cantidad de agua. Están constituidos por un eje central de naturaleza proteica al que están unidos, mediante enlaces covalentes, cadenas de glucosaminoglicanos, que son polisacáridos representados por el condroitín -4 y -6 sulfato en su mayoría, y por el dermatán sulfato y el queratán sulfato. El ácido hialurónico es otro glucosaminoglicano producido por el condrocito que tiene la función de enlazar entre sí los diversos proteoglicanos, formando agregados de proteoglicanos que ocupan los espacios entre las redes de fibrillas colágenas (ver la figura 4). Existen otras proteínas como la fibromodulina, decorina y ancorina, que también tienen sus funciones en la estructura del cartílago. El condrocito está especializado en la biosíntesis y mantenimiento de los componentes de la matriz extracelular a través de la liberación de una serie de moléculas sintetizadas y degradadas durante los procesos anabólicos y catabólicos que, en condiciones normales, se mantienen en equilibrio.

Metabolismo del cartílago en la artrosis

El cartílago articular adulto sano se encuentra en un equilibrio entre los procesos anabólicos (síntesis de matriz extracelular) y catabólicos (degradación de matriz extracelular), en la figura 5 se pueden ver las moléculas que intervienen en estos procesos.

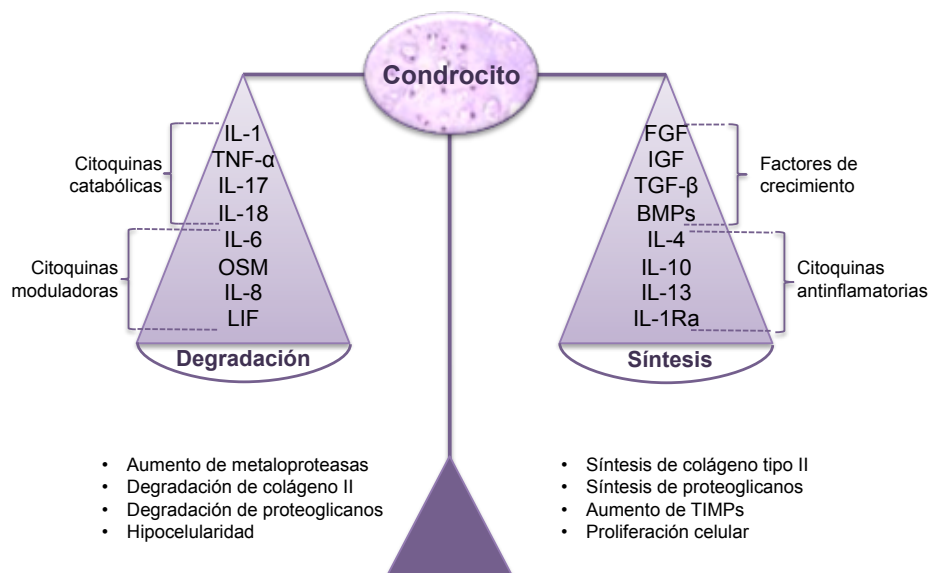


Figura 5. Moléculas que intervienen en el equilibrio entre los procesos anabólicos y catabólicos del cartílago. Las situadas en la parte izquierda promueven la degradación de la matriz extracelular, y las situadas en la parte derecha la síntesis de la matriz extracelular.

Cuando se encuentran en equilibrio estos dos procesos, las moléculas destruidas se reemplazan inmediatamente por nuevas. Sin embargo, durante el envejecimiento y en la artrosis el equilibrio se ve alterado y se produce mayor degradación que síntesis. El origen de este desplazamiento hacia procesos catabólicos todavía no se conoce muy bien, aunque ya se han producido numerosos avances en el estudio de las vías de señalización y transcripción implicadas.

Los procesos anabólicos del cartílago están mediados por factores de crecimiento, citoquinas antiinflamatorias e inhibidores de las proteasas [35-37]. Entre los factores de crecimiento del cartílago, los más conocidos son el factor de crecimiento similar a la insulina (IGF-I), el factor de crecimiento fibroblástico (familia FGF), el factor transformador del crecimiento ($\text{TGF-}\beta$) y las proteínas morfogenéticas del hueso (BMPs). Entre las citoquinas antiinflamatorias destacan las interleucinas 4, 10, 13 y 1Ra (IL-4, IL-10, IL-13, IL-1Ra), que inhiben la actividad y disminuyen la producción de las citoquinas proinflamatorias. Los inhibidores de las proteasas (TIMP) controlan la actividad de las metaloproteasas de la matriz (MMPs), que son las enzimas responsables de la degradación de la matriz del cartílago.

En la artrosis hay una reconocida hipocelularidad respecto al cartílago normal. La regulación de la muerte celular o apoptosis es compleja, puede estar mediada por diversas citoquinas, la familia de proteínas antiapoptóticas Bcl-2 y también intervienen enzimas proteolíticas llamadas caspasas [38]. En cuanto al proceso catabólico en sí, el elemento más importante en la degradación del cartílago articular son las proteasas. Dentro de las proteasas, las que juegan un papel más relevante en la enfermedad son las metaloproteasas (MMPs), que pueden degradar distintos sustratos como colágenos, proteoglicanos, gelatina o fibronectina [39]. También cabe destacar el papel de los mediadores de inflamación como las citoquinas. La interleucina 1β (IL-1 β) es una de las moléculas proinflamatorias más conocida [40-42], es capaz de estimular su propia síntesis y la de otros mediadores como la prostaglandina PGE2 y la enzima ciclooxygenasa 2 (COX-2), así como la de otras interleucinas proinflamatorias como IL-6 e IL-8 y otros mediadores de inflamación como el factor inhibidor de la leucemia (LIF). El factor de necrosis tumoral alfa (TNF- α) también estimula la producción de las

proteasas de matriz. Otras interleucinas que regulan la respuesta catabólica en los condrocitos son las interleucinas 17 y 18 (IL-17, IL-18) [43]. Otra citoquina que parece tener un papel en la regulación de la inflamación es la oncostatina m (OSM) que pertenece al grupo de la IL-6.

1.2.4. El diagnóstico y sus limitaciones

Los principales síntomas de la artrosis son el dolor y la rigidez articular. Las alteraciones radiológicas son la base para el diagnóstico. Sin embargo, es frecuente que pacientes con diagnóstico radiológico de artrosis permanezcan asintomáticos.

El Colegio Americano de Reumatología (ACR) ha establecido los criterios para poder clasificar la artrosis de mano [44], cadera [45] y rodilla [46]. Estos criterios reúnen aspectos clínicos, radiológicos y alteraciones de laboratorio, de manera que se eviten confusiones a la hora de diagnosticar a los pacientes basándose en criterios incompletos.

Artrosis de rodilla
<p>Dolor de rodillas acompañado de, al menos, 4 de estos otros 6 criterios:</p> <ul style="list-style-type: none"> - Edad superior a 50 años - Rigidez matutina de menos de 30 minutos - Crepitación ósea a los movimientos activos - Dolor a la presión sobre los márgenes óseos de la articulación - Hipertrofia articular de consistencia dura - Ausencia de síntomas evidentes de inflamación
Artrosis de mano
<p>Dolor, molestias o rigidez de las articulaciones de las manos, acompañado de:</p> <ul style="list-style-type: none"> - Hipertrofia articular de consistencia dura de 2 o más de 10 articulaciones seleccionadas - Signos de tumefacción presentes en menos de 3 articulaciones metacarpofalángicas - Uno de estos otros 2 criterios: <ul style="list-style-type: none"> - Hipertrofia articular de consistencia dura de 2 o más articulaciones interfalángicas distales (cualquier dedo) - Deformidad de 2 o más de las 10 articulaciones seleccionadas
Artrosis de cadera (criterios clínicos y radiográficos)
<p>Coxalgia presente la mayor parte de los días del mes previo, más 2 de los 3 criterios siguientes:</p> <ul style="list-style-type: none"> - VSG \leq 20 mm/h - Osteofitos femorales o acetabulares en radiografía - Pinzamiento de la interlínea en radiografía

Tabla 3. Criterios clínicos del Colegio Americano de Reumatología para la clasificación de la artrosis [47].

El problema en el diagnóstico de esta enfermedad se presenta en el comienzo de la misma. La mayoría de las personas no tienen síntomas en las fases iniciales, de manera que los pacientes que requieren atención médica se encuentran en fases avanzadas de la enfermedad, cuando ya existen daños irreparables en el cartílago. Por ello, el mayor esfuerzo se está realizando en conseguir un diagnóstico temprano de la artrosis para poder intervenir a tiempo y prevenir la evolución radiográfica. En este sentido, en los últimos años ha surgido un gran interés por encontrar marcadores biológicos específicos de esta patología que faciliten el diagnóstico temprano y mejoren el pronóstico y la evolución de la enfermedad. Existen numerosos estudios de biomarcadores en OA que han aportado una gran cantidad de información [48-50], sin embargo, muy pocos han llegado a la fase de validación. En la tabla 4 están representados los biomarcadores que han llegado a ensayos clínicos hasta hoy [51-53].

Proceso biológico	Biomarcador	Fuente
Metabolismo del cartílago	Telopéptido carboxiterminal del colágeno tipo II (CTX-II)	S/O
	Fragmentos helicoidales del colágeno tipo II (HELIX-II, Coll2-1, Coll 2-1-NO ₂)	S/O
	Epitopos del colágeno tipo II (C2C, C1, 2C)	S
	Keratan sulfato (KS)	S/LS
	Proteína oligomérica de matriz del cartílago (COMP)	S
	Glicoproteína 39 del cartílago (YKL-40)	S
	Pentosidina	S
	NeoEpitopo del colágeno tipo II (TIINE)	O
	C-Propéptidos del colágeno tipo II (CPII, sPIICP)	S
	N-Propéptido del colágeno tipo IIA (PIIANP)	S
	Epitopo 846 del condroitín sulfato (CS846)	S
	Condroitín 4 y 6 sulfato (C6S, C4S)	LS
Metabolismo del hueso	Telopéptido carboxiterminal del colágeno tipo II (CTX-II)	O
	Epitopos del colágeno tipo II (C1, 2C)	S
	Telopéptido aminoterminal del colágeno tipo I (NTX-1)	S/O
	Telopéptido carboxiterminal del colágeno tipo I (CTX-I, ICTP)	S/O
	Telopéptido carboxiterminal del colágeno tipo I (ICTP)	S
	Piridinolina (PYD)	O
	Deoxipiridinolina (DPD)	O
	Sialoproteína ósea (BSP)	S
	Osteocalcina (OC)	S
	Propéptidos carboxiterminal del colágeno tipo I (PICP)	S
	Propéptidos aminoterminal del colágeno tipo I (PINP)	S
Metabolismo de la sinovial	Ácido hialurónico (HA)	S
	Glucosil-galactosil-piridinolina (Glc-Gal-PYD)	S
	Propéptido aminoterminal del colágeno tipo III (PIIINP)	S
Metaloproteasas e inhibidores	Metaloproteasas 1, 3, 8 y 13 (MMPs)	S
	Inhibidor tisular de las metaloproteasas 1,2 (TIMPs)	S
Factores de crecimiento	Factor de crecimiento (TGF-β 1)	S
	Factor de crecimiento insulínico (IGF-1)	S

Tabla 4. Biomarcadores de artrosis empleados en ensayos clínicos, clasificados por el proceso metabólico en el que participan. S: suero; O: orina; LS: líquido sinovial. Tabla adaptada de las referencias [43-45].

A pesar de todos los trabajos realizados en este campo, los marcadores encontrados aún no se pueden emplear a nivel clínico. La solución podría estar en emplear una combinación de varios biomarcadores. Las nuevas tecnologías de análisis a gran escala, proteómica, genómica y metabolómica, pueden aportar gran cantidad de información que conduzca a la identificación de un panel de biomarcadores útil para la práctica clínica.

1.3. Proteómica

La proteómica es la parte de la ciencia que estudia el conjunto de proteínas que se expresan en una célula o tejido bajo unas condiciones determinadas, lo que se conoce con el nombre de proteoma. Este término se empleó por primera vez en 1995 por Wilkins [54] para describir el conjunto de proteínas expresado por un genoma. El estudio de proteomas de forma sistemática ha dado lugar a la aparición de la proteómica.

La proteómica está despertando cada vez un mayor interés. Esto se debe fundamentalmente a que después de haber conseguido la secuenciación del genoma humano es necesario conocer la relación existente entre la expresión de los genes y las funciones biológicas. Esto sólo es posible a través del estudio de las proteínas que codifican, en lo que se conoce como la era postgenómica (figura 6).

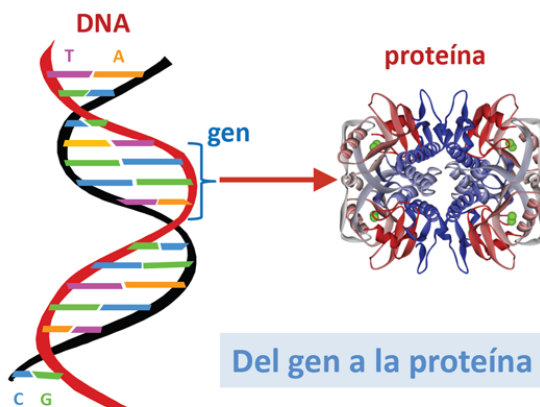


Figura 6. Imagen que representa el paso del genoma al proteoma. Imagen obtenida de la página web <http://guiametabolica.org>.

Mientras que el genoma de un organismo es constante en la totalidad de sus células, el proteoma es un elemento altamente dinámico, cuyos componentes varían como consecuencia de cambios en su entorno, situaciones de estrés, administración de drogas, estado fisiológico o patológico, etc.

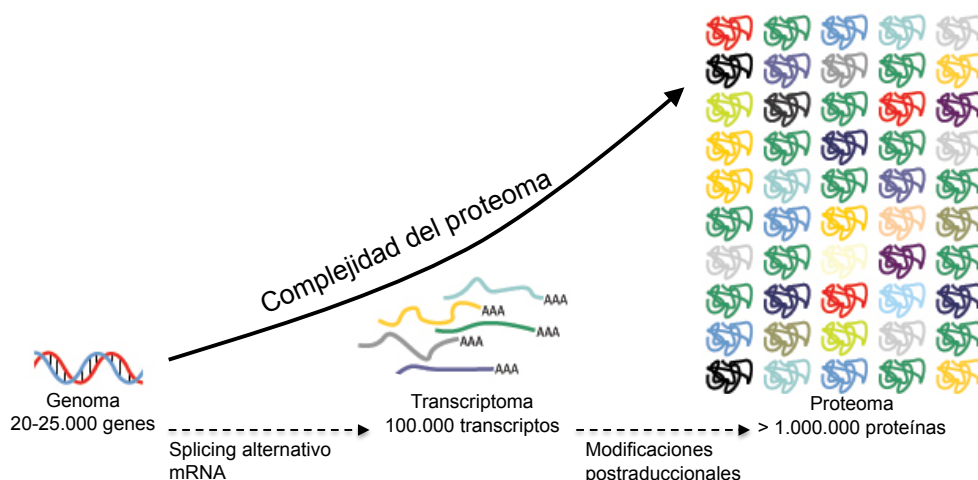


Figura 7. Imagen que representa la complejidad del proteoma en comparación con el genoma, y los factores que influyen en síntesis de las proteínas a partir de los genes. Figura adaptada de la página www.creative-proteomics.com.

Estos factores incrementan de forma considerable la complejidad del proteoma, como consecuencia de la activación o supresión de la expresión de genes, de alteraciones en las interacciones entre las proteínas, o de los cambios producidos por modificaciones postraduccionales [55, 56]. Sin embargo, el dinamismo del proteoma es una ventaja frente al genoma, ya que conseguimos obtener una imagen real de lo que está sucediendo en un organismo. El genoma no tiene en cuenta las posibles modificaciones postraduccionales de las proteínas ni sus interacciones, que son imprescindibles para su función biológica. Tampoco los niveles de expresión genética se corresponden en todos los casos con los niveles de proteína, ya que pueden producirse procesos de transcripción y traducción alternativa, así como mecanismos de degradación de proteínas. En la figura 7 podemos ver cómo afectan estos factores a la complejidad del proteoma.

El gran avance en la proteómica se ha producido gracias al importante desarrollo tecnológico que se ha conseguido en las últimas dos décadas. La primera técnica proteómica surgió en los años 50, la electroforesis de proteínas, que posteriormente evolucionó en el año 1975 a la electroforesis bidimensional en gel (2D-GE). Pero el gran cambio surgió con la aplicación de la espectrometría de masas a la proteómica. Inicialmente, se combinaron las técnicas basadas en electroforesis y la espectrometría de masas: la primera de ellas para separar las proteínas y la segunda para su posterior identificación. Sin embargo, estas primeras estrategias realizaban un análisis poco sensible, excesivamente lento y

costoso. En los últimos años, el gran desarrollo producido en las técnicas de ionización y en los analizadores de masa ha permitido mejorar la sensibilidad y la rapidez en la identificación y secuenciación de proteínas. En los próximos apartados se explicarán con más detenimiento las tecnologías citadas.

Existen numerosas aplicaciones de la proteómica que se pueden englobar en los siguientes grupos (ver figura 8):

1. Caracterización del proteoma de un organismo. El propósito de este estudio es la identificación del mayor número de proteínas expresadas por un organismo en una condición biológica particular.
2. Interacción de proteínas. En este tipo de análisis se estudian las interacciones entre proteínas y también entre proteína y otra molécula. Estos resultados contribuyen significativamente al establecimiento de mapas de interacciones y al esclarecimiento del mecanismo de determinadas funciones biológicas.
3. Modificaciones post-traduccionales. Las modificaciones post-traduccionales se producen en las proteínas una vez sintetizadas, inducen una gran variedad de cambios estructurales y funcionales y constituyen un mecanismo esencial de regulación de su actividad biológica. El conocimiento de estas modificaciones es clave para comprender la biología celular, el mecanismo de acción de muchos fármacos y para diseñar nuevas estrategias terapéuticas.
4. Proteómica comparativa. Aborda la identificación de proteínas que sufren alteraciones en sus niveles de expresión a consecuencia de cambios en su entorno, situaciones de estrés, administración de drogas, efectores o señales bioquímicas, su estado fisiológico o su estado patológico. Es lo que se conoce como proteómica de expresión diferencial. Esta aplicación es la más importante en ciencias biomédicas, ya que permite determinar cuáles son las proteínas que intervienen en estos procesos, llegando a caracterizar un mecanismo molecular o identificar dianas farmacológicas. Además posibilita identificar proteínas cuyos cambios de expresión resulten característicos del proceso, obteniendo así posibles marcadores biológicos.

En este trabajo nos hemos centrado en la proteómica comparativa o diferencial, ya que la finalidad del mismo es la búsqueda de biomarcadores para enfermedades reumáticas.

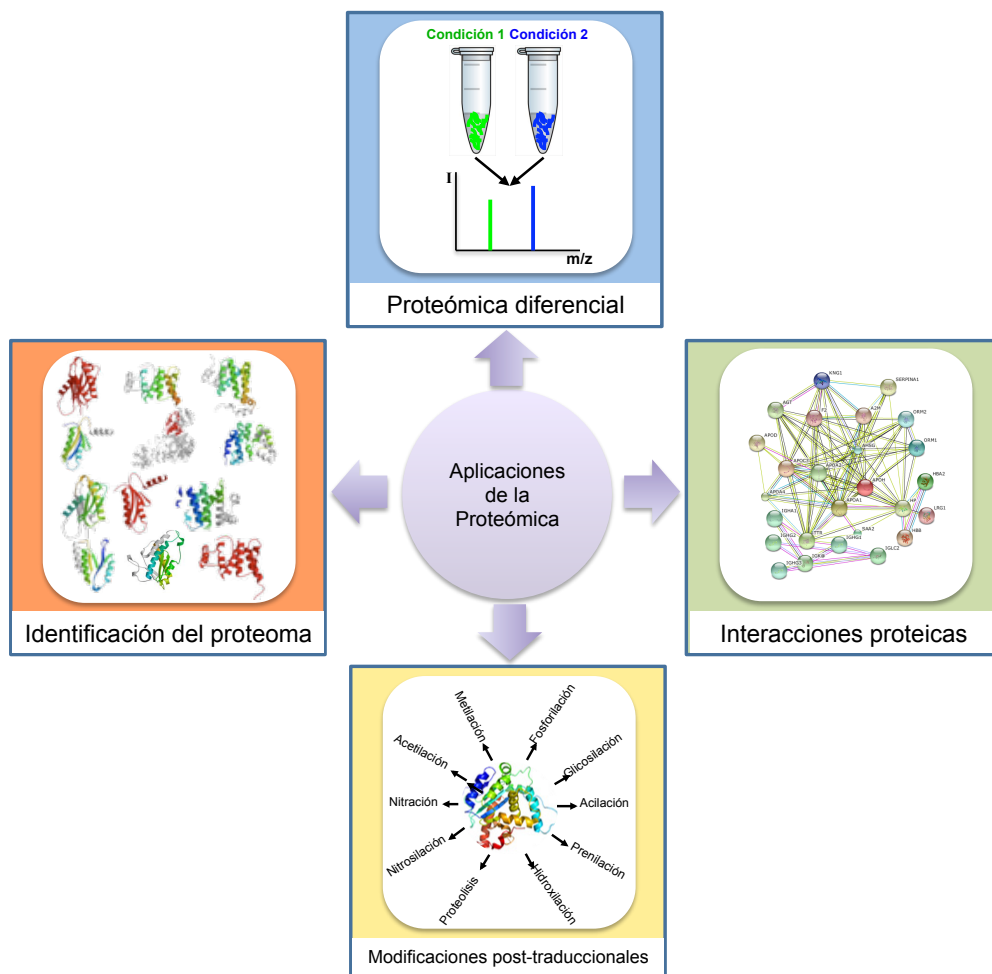


Figura 8. Principales aplicaciones de la Proteómica.

1.3.1. Búsqueda de biomarcadores

Un biomarcador es una característica que se puede medir y evaluar objetivamente como indicador de un proceso biológico, proceso patogénico o de respuesta farmacológica a una intervención terapéutica [57].

El término biomarcador abarca proteínas, fragmentos proteicos, metabolitos, carbohidratos, ácidos nucleicos (ADN y ARN), características celulares e imágenes [58]. Recientemente se ha establecido una clasificación de los biomarcadores en húmedos y secos. Dentro de los marcadores húmedos se encuentran las sustancias que proceden de los fluidos o tejidos humanos (suero, plasma, orina, etc), mientras que dentro de los marcadores secos se encuentran los cuestionarios clínicos y las imágenes (resonancia magnética, radiografías, ultrasonidos, etc.) [51, 52, 59].

Los biomarcadores se pueden emplear en clínica en el diagnóstico, en la monitorización de una enfermedad, como guía en terapias dirigidas o para evaluar respuestas terapéuticas [60]. En el caso de la artrosis, tal y como se explicó en el apartado del diagnóstico y sus limitaciones, se intenta en la actualidad buscar biomarcadores que faciliten el diagnóstico temprano de la enfermedad, que mejoren su pronóstico y que permitan optimizar tratamientos individualizados. En este sentido, la proteómica tiene un papel importante, ya que las proteínas tienen un gran potencial biomarcador. Además, las técnicas proteómicas son de alto rendimiento y permiten identificar un gran número de proteínas en un único experimento.

Las técnicas proteómicas se basan principalmente en la separación de un gran número de proteínas y en su posterior identificación mediante espectrometría de masas (MS). Las técnicas de separación de proteínas pueden ser de dos tipos, las basadas en electroforesis en gel y las basadas en cromatografía líquida, conocidas también como técnicas sin gel (gel-free). Cada una tiene sus ventajas e inconvenientes. El escoger una técnica u otra depende fundamentalmente del tipo de muestra. Esta tesis se realizó únicamente con muestras de suero humano, que poseen unas características particulares que describiremos a continuación.

Muestras de suero humano

Las muestras de plasma o suero son las más empleadas en búsqueda de biomarcadores por su fácil obtención, al no ser necesario recurrir a técnicas invasivas para conseguir la muestra tal y como ocurre con las muestras de tejidos.

Sin embargo el suero tiene sus limitaciones para los estudios proteómicos. Se trata de muestras especialmente complejas debido a la gran cantidad de proteínas presentes en las mismas, estimadas en más de diez mil proteínas diferentes [61]. Además, estas proteínas se encuentran en concentraciones muy dispares, dando lugar a un elevado rango dinámico que abarca más de 10 órdenes de magnitud [62, 63]. En la figura 9 se representa dicho rango dinámico, que permite observar tres grandes grupos: las proteínas clásicas del plasma que se encuentran a elevadas concentraciones (mg/ml), las proteínas que provienen de tejidos degradados con menor concentración que las anteriores ($\mu\text{g/ml}$), y las citoquinas e interleucinas que se encuentran a concentraciones muy bajas (pg/ml). Aparte del problema del elevado rango dinámico, es importante tener en cuenta el papel de determinadas proteínas dentro del proteoma del suero, como el grupo de las proteínas más abundantes que suelen denominarse las Top 20. Se trata de un grupo de 22 proteínas que representan en torno al 99 % de la cantidad total proteica presente en el suero [64].

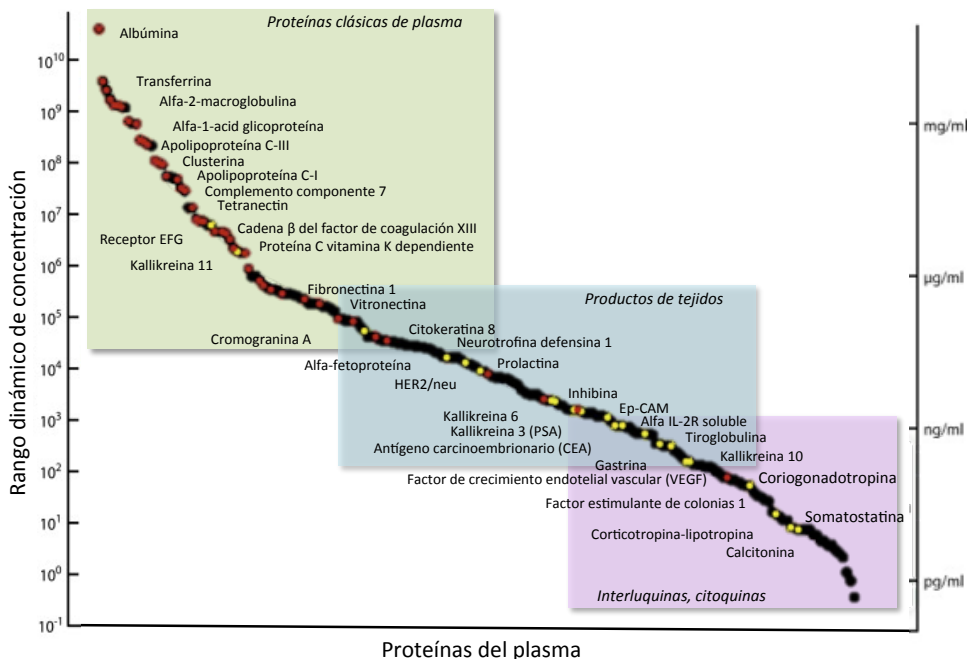


Figura 9. Rango dinámico de concentración de las proteínas del plasma. En círculos amarillos están representados los biomarcadores empleados en clínica y en círculos rojos se encuentran las proteínas descubiertas por el Proyecto Proteoma del Plasma de la HUPO (Organización del Proteoma Humano). EFG: factor de crecimiento epidérmico; EpCAM: molécula de adhesión de células epiteliales. Figura adaptada de Schiess R. y colaboradores [54].

En este grupo destaca la albúmina, que representa más del 50 % de la carga proteica. En la figura 10 se puede ver la distribución de las proteínas más abundantes del suero y el papel que desempeña la albúmina.

Normalmente, este grupo de proteínas no contiene información relevante para ejercer el papel de biomarcadores, y sin embargo son tan abundantes que no permiten la identificación de otras proteínas con mayor potencial. Como consecuencia, es necesario aplicar métodos para reducir la presencia de dichas proteínas en el suero, o incluso llegar a eliminarlas. Estos métodos se conocen con el nombre de métodos de depleción [65-68]. En la sección 1.3.3. se explican en detalle los métodos de depleción generalmente empleados en proteómica.

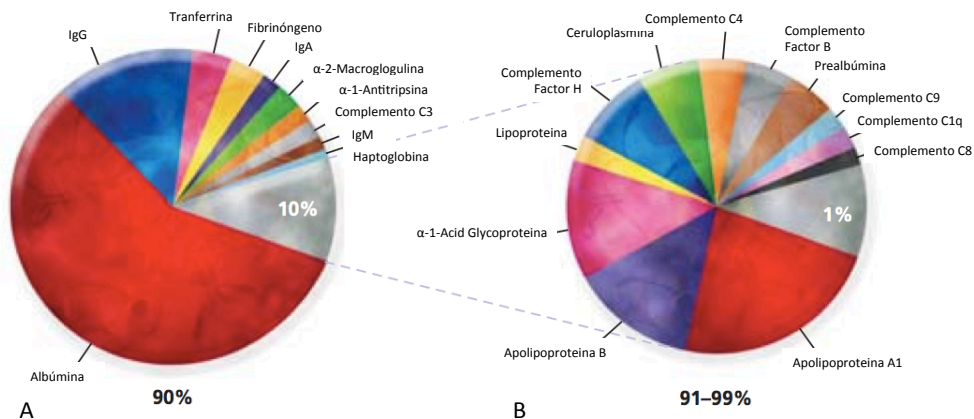


Figura 10. Distribución de las proteínas más abundantes del plasma. En la figura A se encuentran las 10 proteínas más abundantes, que representan entorno al 90 % de la masa total, y el 10% restante está representado en la figura B, en donde se suman 12 proteínas más al grupo de las abundantes. Figura adaptada de SigmaAldrich.

1.3.2. Estrategias proteómicas para la búsqueda de biomarcadores

Las estrategias proteómicas para la búsqueda de biomarcadores son numerosas [69-72], aunque como comentamos anteriormente todas constan de una primera fase de separación de las proteínas (y/o péptidos) seguida de una segunda fase de identificación de las mismas mediante espectrometría de masas.

Como se describe en el apartado anterior, la muestra de suero es un tipo de muestra extremadamente compleja debido al elevado número de proteínas presentes y a su rango dinámico, lo que hace necesario dividir el proteoma total del suero en fracciones más sencillas de analizar. Por ello, uno de los pasos más importantes es el tipo de separación que se va a aplicar. Se pueden escoger técnicas de electroforesis en gel o técnicas libres de gel (generalmente métodos cromatográficos) [73, 74]. Además se debe incluir un paso previo de depleción para evitar el problema de las proteínas abundantes del suero. En la figura 11 están representadas las estrategias proteómicas más empleadas en la búsqueda de biomarcadores en muestras de suero.

Las estrategias basadas en gel tienen una serie de ventajas y desventajas en comparación con las separaciones cromatográficas. La ventaja más destacada es la capacidad para detectar proteínas con modificaciones post-traduccionales con alta resolución. Las desventajas principales son la dificultad para separar determinados tipos de proteínas de tamaño y acidez o basicidad extremos, la limitación en el rango dinámico de las diferentes tinciones que se aplican y la baja capacidad para analizar un gran número de muestras a la vez.

Las estrategias de separación no basadas en gel suelen ser métodos de separación por cromatografía líquida (LC). Las principales ventajas de estos métodos es que son más fáciles de automatizar, permitiendo el análisis de un mayor número de muestras, y tienen mayor capacidad de detectar proteínas menos abundantes del suero [75-77].

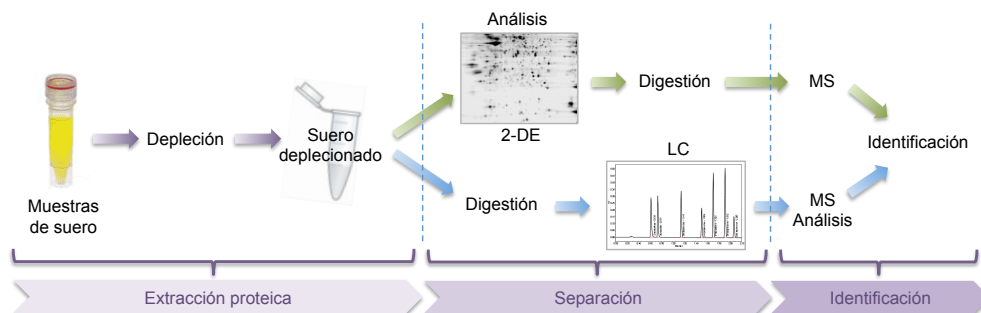


Figura 11. Estrategias proteómicas empleadas en la búsqueda de biomarcadores en muestras de suero, basadas en la separación proteica por electroforesis en gel o por cromatografía líquida. 2-DE: electroforesis bidimensional en gel; LC: cromatografía líquida; MS: espectrometría de masas. Figura adaptada de Ruiz-Romero y colaboradores [64].

1.3.3. Métodos de depleción

Existen diferentes métodos para deplecionar las proteínas abundantes del suero, como la cromatografía de inmunoafinidad, columnas de extracción en fase sólida, reacciones químicas específicas o kits comerciales diseñados para abarcar el rango dinámico de concentración de las proteínas [65-68, 78-85].

Los métodos más empleados se basan en el empleo de anticuerpos específicos frente a las proteínas más abundantes. Estos anticuerpos se pueden inmovilizar en columnas cromatográficas (cromatografía de inmunoafinidad) o en cartuchos comerciales de extracción en fase sólida, y pueden reconocer una o varias de las proteínas abundantes. La inmunodepleción más eficaz disponible comercialmente es la que emplea columnas para cromatografía líquida de alta eficacia con anticuerpos inmovilizados frente a las 20 proteínas más abundantes. En la figura 12 se representa el principio de la depleción por inmunoafinidad.

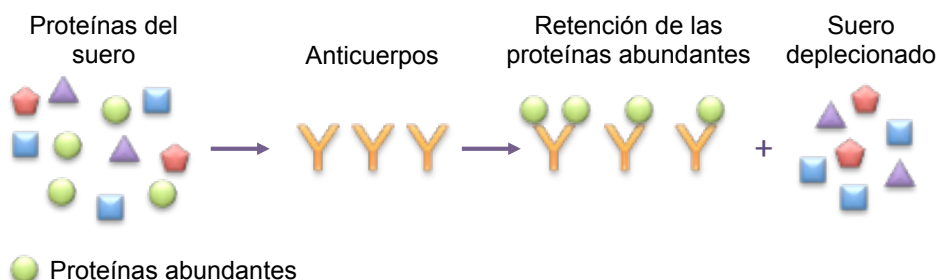


Figura 12. Método de depleción basado en inmunoafinidad. El suero se pasa por un dispositivo en donde se encuentran los anticuerpos específicos de las proteínas abundantes, de modo que estas proteínas se quedan retenidas en el dispositivo, mientras que el suero deplecionado se recoge en la fracción no retenida.

La desventaja de aplicar cualquier método de depleción es que algunas proteínas de interés pueden ser deplecionadas junto con las abundantes. Este hecho es casi imposible de controlar debido a las interacciones que se establecen entre ambos tipos de proteínas. Se ha comprobado que uno de los métodos que mejor evita esta depleción no deseada es la cromatografía líquida de inmunoafinidad [86], por tratarse de una unión específica de las proteínas abundantes con sus anticuerpos.

1.3.4. Métodos de digestión de proteínas

Existen diversos métodos para romper las proteínas dando lugar a sus fragmentos, llamados péptidos. Los enzimas denominados peptidasas catalizan la rotura hidrolítica de los enlaces peptídicos. Algunas peptidasas cortan específicamente en el enlace peptídico adyacente a un aminoácido determinado, dando lugar a una digestión proteica predecible y de forma reproducible. Este es el caso de uno de los enzimas más empleados en la proteómica, la tripsina. Este enzima hidroliza un enlace peptídico después de los residuos de lisina (Lys) y arginina (Arg), con la excepción de que la Lys o Arg estén N-ligadas a un ácido aspártico (Asp) o en algunos casos en los que una prolina se encuentre antes de la Lys o Arg. En la tabla 5 se pueden ver algunos de los enzimas más empleados y sus sitios de corte.

En general, los protocolos de digestión necesitan un paso previo de rotura de los puentes disulfuro de las proteínas, denominado reducción, seguido de la

alquilación de los grupos sulfhidrilo para evitar la formación de los puentes disulfuro de nuevo.

El proceso de digestión completo suele ser una parte de la preparación de la muestra que consume mucho tiempo, en torno a 12-16 h. Se han probado diferentes métodos para acelerar este proceso y reducir el tiempo de tratamiento de muestra, como aplicar temperatura, energía de microondas, energía de ultrasonidos (US), alta presión, energía de infrarrojos (IR) o micro-reactores de

Proteasas	Puntos de rotura
Tripsina	Lys, Arg (C)
Quimotripsina	Phe, Trp, Tyr (C)
Asp-N-proteasa	Asp, Glu (N)
Pepsina	Phe, Trp, Tyr (N)
Endopeptidasa Lys C	Lys (C)

Tabla 5. Proteasas empleadas en la digestión de proteínas. Lys: lisina; Arg: arginina; Phe: fenilalanina; Trp: triptófano; Tyr: tirosina; Asp: ácido aspártico; Glu: ácido glutámico; C: extremo carbonilo; N: extremo amino.

tripsina inmovilizada [87, 88]. La temperatura que se emplea habitualmente es de 37°C para realizar la digestión durante la noche, pero se ha realizado algún estudio en el que se comprueba que a 50-60°C se reduce la autólisis del propio enzima y se disminuye el tiempo de digestión [89]. No obstante, no hay muchos estudios en donde se compruebe una mejoría evidente. En cuanto a la digestión asistida por microondas, se han realizado diversos experimentos a diferentes temperaturas en los que a determinadas condiciones se acelera el proceso de digestión [90-93], pero parece que el efecto del calor sobre la propia reacción no es beneficioso [88]. Wang y colaboradores emplearon en 2008 la energía de los infrarrojos para acelerar la digestión proteica [94-96], consiguiendo reducir a 5 minutos el proceso, aunque diferentes estudios posteriores no encontraron muchas ventajas respecto al método clásico de 12h a 37°C [88] y es necesario disponer de un sistema de IR adaptado. El empleo de ultrasonidos también implica disponer de un equipo específico, pero consigue reducir el tiempo de la reacción a 5 min, por lo que parece ser uno de los métodos más prometedores para acelerar la digestión [97-99]. También existe tripsina inmovilizada en diferentes soportes sólidos, que se conocen como micro-reactores. Estos sistemas se están convirtiendo en los más empleados para acelerar la digestión, ya que el enzima se puede reutilizar, tiene un bajo consumo de reactivos y evita los problemas derivados de la autólisis del propio enzima [100-104].

1.3.5. Métodos de separación basados en gel

Uno de los primeros métodos de separación de proteínas fue la electroforesis en geles de dodecilsulfato sódico (SDS) y poliacrilamida (PAGE), en los que las proteínas cargadas migran según su peso molecular bajo la influencia de un campo

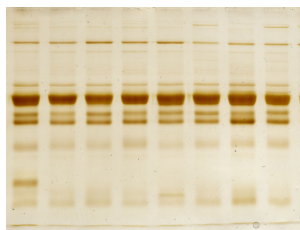


Figura 13. Gel del 10 % de acrilamida teñido con plata, en el que se separaron las proteínas de diferentes muestras de suero.

eléctrico. Los geles de poliacrilamida están formados por la polimerización de la acrilamida con la bisacrilamida, creando una red porosa a través de la cual las proteínas migran en función de la relación carga/masa. El SDS se emplea para proporcionar una carga neta negativa a todo el entorno, de manera que la carga intrínseca de las proteínas se vuelve insignificante y la migración sólo depende de su masa. En la figura 13 se muestra la imagen de un gel

de acrilamida realizado en nuestro laboratorio con muestras de suero y teñido con plata.

La electroforesis clásica evolucionó posteriormente a la electroforesis bidimensional (2-DE), que consta de dos separaciones, con el fin de obtener una mejor resolución de las proteínas. En la primera etapa, llamada isoelectroenfoque, las proteínas se separan en función de su punto isoeléctrico (pI) bajo un gradiente de pH, y a continuación se realiza una segunda separación en geles de SDS-PAGE en función de su peso molecular. Esta técnica permite la resolución de muestras proteicas complejas y es más sensible que la anterior [105], al separar proteínas del mismo peso molecular que tengan distinto punto isoeléctrico. En la figura 14 está representado el proceso de la electroforesis bidimensional de forma esquemática. Esta técnica ha sido una de las más empleadas en las últimas décadas para la separación de muestras complejas, es la única que proporciona un mapa de proteínas intactas que reflejan cambios al nivel de expresión de proteínas, isoformas o modificaciones post-traduccionales. La electroforesis bidimensional fue introducida por primera vez por O'Farrell en el año 1975 [106].

La técnica de 2-DE tiene sus limitaciones. Entre otras, la reproducibilidad entre geles es muy complicada de conseguir [107], no es suficientemente sensible para detectar proteínas poco abundantes, y su rango dinámico es limitado. Debido a

ello, hace unos años Ünlü y colaboradores desarrollaron la electroforesis diferencial en gel (DIGE) [108]. En esta técnica las muestras proteicas que se quieren comparar se marcan previamente con moléculas fluorescentes

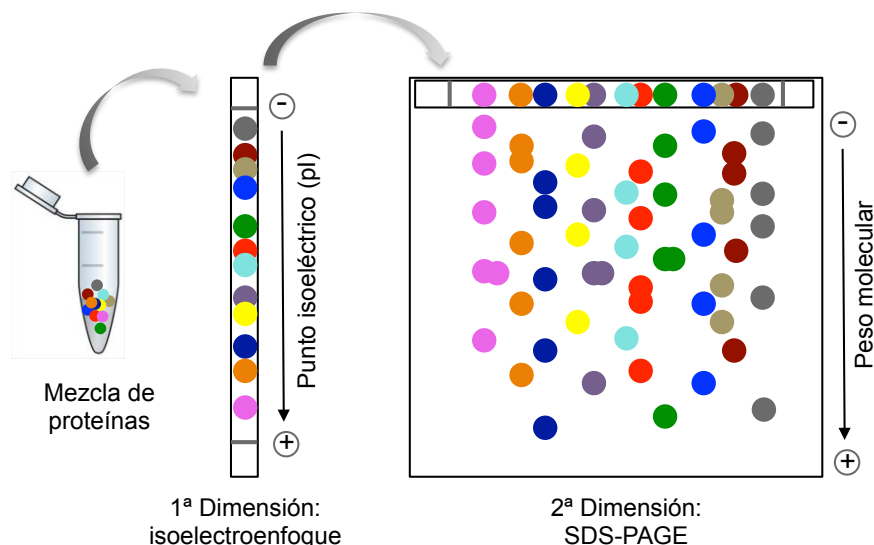


Figura 14. Proceso de una separación de proteínas mediante la técnica de electroforesis bidimensional. SDS-PAGE: dodecilsulfato sódico-electroforesis en geles de poliácridamida.

denominadas CyDyes (Cy2, Cy3 y Cy5). Estas moléculas son cianinas que tienen la misma masa y carga, y que cuando se excitan emiten a diferentes longitudes de onda, permitiendo su detección de forma independiente. Son muy sensibles, de modo que proporcionan un rango dinámico muy amplio para la cuantificación. Normalmente Cy3 y Cy5 se emplean para marcar dos condiciones y el tercer fluoróforo (Cy2) se utiliza para marcar un estándar interno constituido por la mezcla a partes iguales de todas las muestras a analizar. Este estándar se emplea en todos los geles del experimento, de esta forma se mejora la reproducibilidad del método, ya que las dos condiciones se corren a la vez en un mismo gel y el estándar interno ayuda a normalizar las diferencias técnicas que se puedan producir al manipular las muestras y entre los distintos geles que se empleen. En la figura 15 se puede ver de forma sencilla en qué consiste la técnica DIGE.

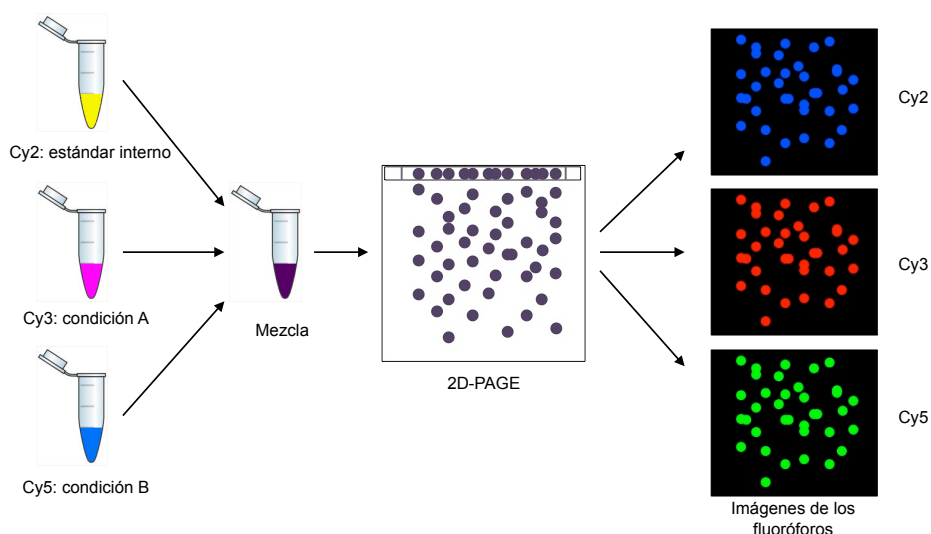


Figura 15. Esquema representativo de la técnica de electroforesis diferencial en gel (2D-DIGE) para la separación y cuantificación de proteínas. Las dos muestras (condición A y condición B) se marcan con los fluoróforos Cy3 y Cy5 respectivamente, y el estándar interno formado por la mezcla de cantidades iguales de las dos condiciones se marca con el Cy2. Se mezclan las tres muestras marcadas y se separan mediante electroforesis bidimensional. Por último, los geles se escanean a diferentes longitudes de onda, obteniendo las imágenes de los tres fluoróforos.

1.3.6. Métodos de separación cromatográficos

El tipo de cromatografía más empleada en proteómica es la cromatografía de líquidos, ya que puede ser acoplada directamente al espectrómetro de masas y se puede aplicar tanto a proteínas como a péptidos. Los métodos de cromatografía de líquidos están basados en el reparto de las proteínas o péptidos entre dos fases, la fase móvil y la fase estacionaria (columna cromatográfica). Dicho reparto se puede producir por diferentes tipos de interacciones entre la muestra y las dos fases, interacciones electrostáticas, hidrofóbicas, afinidad, tamaño, etc. De acuerdo al tipo de interacciones que se producen existen diversos tipos de cromatografía líquida. Los más empleados en proteómica son: fase reversa (RP), intercambio iónico (IEX), exclusión por tamaño, interacciones hidrofílicas (HILIC) y afinidad [71, 109-111]. En la figura 16 se representan los distintos tipos de cromatografía empleados en proteómica.

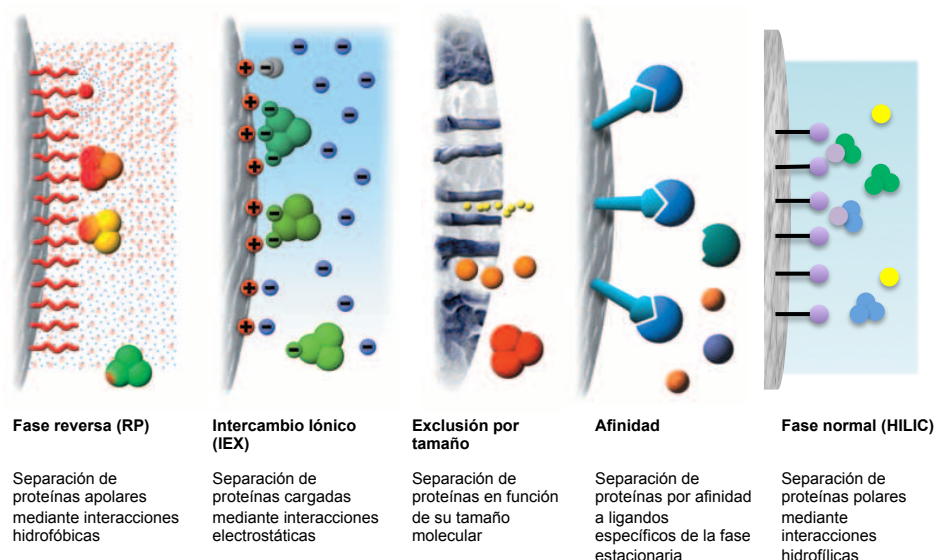


Figura 16. Tipos de cromatografía de líquidos más empleados en la separación de proteínas. Figura adaptada de GE Healthcare Life Sciences.

En cualquier tipo de cromatografía, los componentes se distribuyen de modo distinto entre la fase móvil y la fase estacionaria. Los fuertemente retenidos por la fase estacionaria se mueven lentamente con el flujo de la fase móvil, mientras que los componentes que se unen débilmente a la fase estacionaria se mueven con rapidez. Como resultado de la distinta movilidad, los componentes de la muestra se separan en fracciones discretas que pueden analizarse por separado. Normalmente las técnicas de separación cromatográficas se realizan después de llevar a cabo la digestión de las proteínas.

Cromatografía de fase reversa

La cromatografía de fase reversa separa las proteínas por su hidrofobicidad. Se trata de una cromatografía de reparto en la que la fase estacionaria es apolar y la fase móvil polar. Las fases estacionarias más empleadas son los hidrocarburos de cadena larga (C_{18}), mientras que la fase móvil es de carácter orgánico (acetonitrilo, ACN). Debido al carácter apolar de los péptidos, éstos se quedan retenidos en la columna. La elución gradual se realiza incrementando el % de fase orgánica, de manera que primero eluyen los péptidos más polares y por último los menos polares. Se trata de la cromatografía más empleada en proteómica debido a su alta resolución, y a la posibilidad de ser acoplada directamente con el espectrómetro de

masas, ya que los tampones que emplea son compatibles con la ionización por electrospray (ESI).

Cromatografía de intercambio iónico (IEX)

En la cromatografía de intercambio iónico las proteínas se separan en función de su carga. Los analitos se unen al soporte cromatográfico a través de interacciones electrostáticas entre las cargas eléctricas presentes en las proteínas y las cargas opuestas presentes en la fase estacionaria. En este caso, las proteínas cargadas con carga opuesta a la del soporte se quedarán retenidas, y la elución se realiza aumentando la fuerza iónica de la fase móvil, de forma que primero eluyen las proteínas menos cargadas y por último las de mayor carga. Este tipo de cromatografía puede ser de dos tipos: de intercambio catiónico (fase estacionaria cargada negativamente) o de intercambio aniónico (fase estacionaria cargada positivamente).

Cromatografía de exclusión por tamaño

Esta cromatografía también se conoce por filtración en gel o de penetrabilidad sobre geles. La cromatografía de exclusión por tamaño separa a las proteínas en base a su tamaño molecular, no se producen interacciones químicas o físicas entre los analitos y la fase estacionaria. Los rellenos para la fase estacionaria están constituidos por pequeñas partículas que forman una red de poros uniforme en los que pueden difundir las moléculas del soluto y del disolvente. Las moléculas son atrapadas eficazmente en los poros, y el tiempo de residencia medio en los poros depende del tamaño de la molécula. Las proteínas más grandes que el tamaño medio de los poros del relleno son excluidas, de forma que son las primeras que eluyen (no se retienen). En cambio, las proteínas de menor tamaño pueden penetrar a través del laberinto de poros quedando más tiempo atrapadas, y por tanto, eluyendo las últimas.

Cromatografía de afinidad

La cromatografía de afinidad separa las proteínas en base a las interacciones entre las proteínas y los ligandos específicos de la fase estacionaria. Se trata de una

cromatografía altamente selectiva y de gran resolución. Los ligandos del soporte cromatográfico pueden ser de diversos tipos: anticuerpos, enzimas, metales, etc. Cuando la mezcla de proteínas entra en contacto con la fase estacionaria, se retienen sólo aquellas proteínas que se pueden unir al ligando específico, y la elución se puede realizar por desestabilización de la unión ligando-proteína o por la adición de un ligando competitivo.

Cromatografía de interacciones hidrofílicas o de fase normal (HILIC)

La cromatografía de fase normal se emplea para separar proteínas polares en base a las interacciones hidrofílicas entre el analito y la fase estacionaria. En este tipo de cromatografía la fase estacionaria es un soporte polar, mientras que la fase móvil es apolar. Este tipo de separación se suele emplear para separar aquellas proteínas o péptidos que no se pueden separar por cromatografía de fase reversa.

Cromatografía multidimensional

El proteoma del suero es una mezcla muy compleja de proteínas, por lo que para su análisis por espectrometría de masas es necesario incrementar la separación de los péptidos empleando varios métodos cromatográficos [112, 113]. Es lo que se conoce por cromatografía multidimensional. Una de las estrategias más empleada en proteómica consta de una primera separación por cromatografía de intercambio iónico seguida de una separación por cromatografía de fase reversa. Sin embargo, en los últimos años han ido surgiendo diferentes combinaciones, en las que los distintos métodos cromatográficos se acoplan a una separación final de fase reversa a bajo pH. Entre ellos está la cromatografía de fase reversa a alto pH, la de exclusión por tamaño o la cromatografía de interacciones hidrofílicas. La ventaja de utilizar la cromatografía de fase reversa a bajo pH como última etapa es que permite acoplarla al espectrómetro de masas directamente, gracias a la compatibilidad de la fase móvil con ESI, a su alta resolución y a su capacidad de eliminar sales de las muestras, haciendo más sencillo el análisis mediante espectrometría de masas.

1.3.7. Espectrometría de masas

La espectrometría de masas es un método analítico que determina la relación masa/carga (m/z) de átomos o moléculas. Este tipo de análisis se compone de varias etapas: la primera es la ionización de la muestra, la segunda consiste en la separación de los iones según su relación m/z y la última etapa es el recuento de cada tipo de iones mediante un detector adecuado.

La espectrometría de masas tuvo su origen en 1886, cuando Goldstein descubrió los iones positivos en un tubo de descarga eléctrica a baja presión. Posteriormente, Wien descubrió que los campos eléctricos o magnéticos fuertes refractaban los iones, y en 1899 construyó un dispositivo con campos eléctricos y magnéticos paralelos que separaban los rayos positivos según su relación carga/masa. Casi a la vez, en 1912, Joseph Thomson continuó con los estudios de los rayos catódicos de Goldstein, y al igual que Wien pero de diferente forma consiguió crear el primer espectrómetro de masas. Las técnicas más modernas aparecieron en los años 1950-1960, cuando Dehmelt y Paul desarrollaron la trampa de iones. A finales de los 80 se desarrollaron las dos técnicas de ionización más empleadas hoy en día, Fenn [114] creó la ionización por electrospray, Tanaka [115] y Karas [116] desarrollaron de forma independiente y casi a la vez, la ionización/desorción por láser asistido por matriz (MALDI).

Los espectrómetros de masas están formados básicamente por una fuente de iones, un analizador y un detector. En la figura 17 se representan los componentes principales de cualquier espectrómetro de masas. Estos espectrómetros requieren un complejo sistema de vacío para mantener una baja presión en todos los componentes. El sistema de entrada introduce la muestra en la fuente de iones, donde los componentes se convierten en iones gaseosos. Este flujo de iones gaseosos es acelerado en el analizador de masas, donde se separan en función de su relación m/z , y por último llegan al detector, que convierte el haz de iones en una señal eléctrica que pueda ser procesada, almacenada e interpretada, dando lugar al espectro de masas.

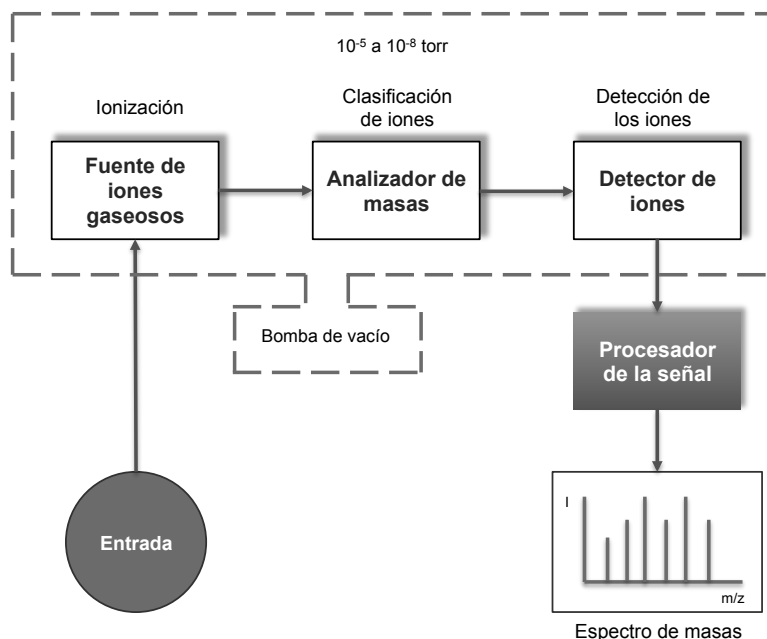


Figura 17. Componentes de un espectrómetro de masas.

1.3.7.1. Fuentes de iones

La fuente de iones es la encargada de convertir los analitos en iones gaseosos (ionización). El aspecto de los espectros de masas depende en gran medida del método utilizado para la formación de los iones. Las fuentes más empleadas en proteómica son la de ionización por electronebulización o electrospray (ESI) y la fuente de desorción/ionización por láser asistida por una matriz (MALDI). Ambas fuentes son fuentes de desorción, en donde las muestras en estado sólido o líquido se transforman directamente en iones gaseosos.

Ionización por electronebulización (ESI)

La ionización por electronebulización o electrospray se aplica a muestras líquidas, por lo que resulta fácil acoplarla a la cromatografía de líquidos. La ionización se realiza en condiciones atmosféricas de presión y temperatura. En la figura 18 se muestra el sistema de ionización ESI. La muestra líquida pasa a través de una aguja capilar de acero sometida a alto voltaje (2-6 kV), que produce la formación de un espray cargado (cono de Taylor). Las gotas iniciales se van reduciendo de

tamaño por evaporación, generándose una mayor densidad de carga que provoca una fuerza de repulsión de Coulomb que puede llegar a superar la tensión superficial de las gotas, causando la rotura de las mismas (explosión de Coulomb) y formando gotas más pequeñas. Este proceso continúa hasta que las gotas son lo suficientemente pequeñas para producir la desorción de los iones, que alcanzan la fase gaseosa. Se suele emplear un capilar caliente o un gas de secado (normalmente nitrógeno, N_2) para ayudar a la formación del electrospray. Debido a que el proceso de desorción se produce a presión atmosférica, los iones gaseosos son atraídos dentro del espectrómetro por la diferencia de presión. Con este tipo de ionización se forman iones de carga múltiple, generalmente con

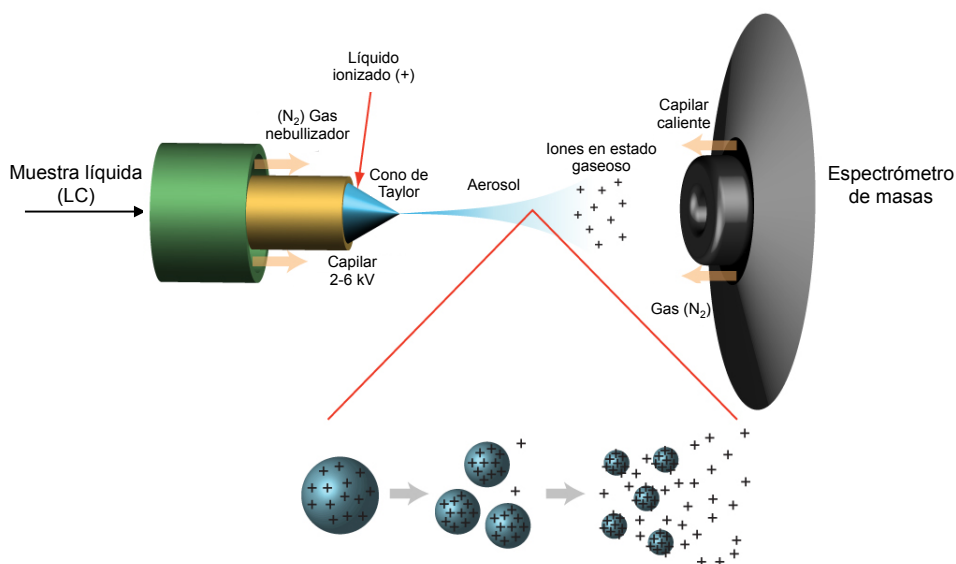


Figura 18. Esquema de la ionización por electrospray. Figura adaptada de <http://www.lamondlab.com/MSResource/LCMS/MassSpectrometry/electrosprayionisation.php>.

carga positiva, aunque se puede trabajar con iones de carga negativa como en los péptidos fosforilados y sulfatados.

Desorción/ionización por láser asistida por una matriz (MALDI)

En esta técnica de ionización las muestras se encuentran en estado sólido. La muestra se mezcla con una matriz orgánica y se deja cristalizar al aire sobre una superficie metálica, de modo que la ionización se produce por sublimación al incidir un láser sobre ella.

La matriz tiene la característica de absorber la energía del láser, debido a que se trata de una molécula orgánica conjugada. Los grupos aromáticos de la matriz absorben energía a la longitud de onda del láser (generalmente láser pulsado de N₂ o Nd:YAG), causando la ionización de la matriz seguida por la transferencia de energía y protones al analito. En la figura 19 se puede ver el proceso de ionización por MALDI. Las matrices más empleadas en proteómica son el ácido α -ciano-4-

hidroxicinámico (CHCA), el ácido 2,5-dihidroxibenzoico (DHB) y el ácido 3,5-dimetoxi-4-hidroxicinámico (ácido sinapínico), que absorben a una longitud de onda similar (en torno a 337 nm). En la tabla 6 se puede ver la estructura y la longitud de onda de las principales matrices empleadas en proteómica. La preparación de la matriz es una parte fundamental de este tipo de ionización, y tiene un efecto directo en la calidad de los espectros de masas. Se han desarrollado numerosos protocolos de preparación en los que la matriz se disuelve en soluciones volátiles (como el acetonitrilo), de modo que al evaporarse más rápido se consigue una capa de cristales más homogénea. Cuando la cristalización no es homogénea, disminuye la resolución, la reproducibilidad y se obtiene una baja correlación entre la intensidad y la concentración del analito. En MALDI, la mayoría de los péptidos y proteínas se ionizan con una carga positiva.


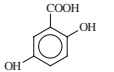
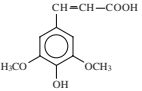
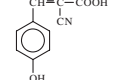
Matriz	Estructura	Longitud de onda
Ácido nicotínico		UV 266nm
Ácido 2,5-dihidroxibenzoico		UV 337 y 353nm
Ácido sinapínico		UV 337 y 353nm
Ácido α -ciano-4-hidroxicinámico		UV 337 y 353nm

Tabla 6. Principales matrices para MALDI empleadas en proteómica. UV: ultravioleta.

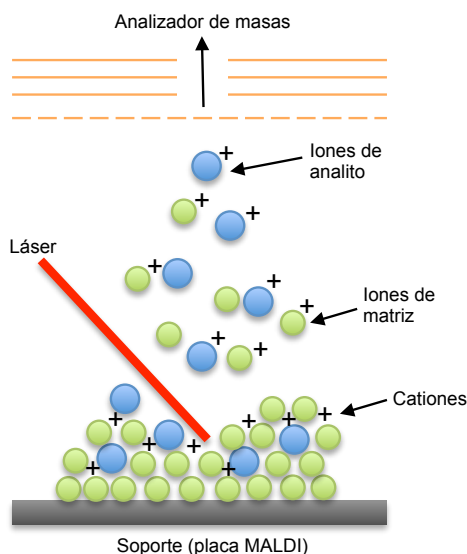


Figura 19. Desorción/ionización por láser asistida por matriz (MALDI).

1.3.7.2. Analizadores de masas

Los analizadores de masas separan los iones en base a su relación m/z . Lo ideal es que el analizador sea capaz de distinguir entre pequeñas diferencias de masa. Los más empleados en proteómica son el de tiempo de vuelo (TOF), el triple cuadrupolo, la trampa de iones, de transformada de Fourier (FT) y el Orbitrap. En la figura 20 se muestra un esquema de los principales analizadores de masas [77].

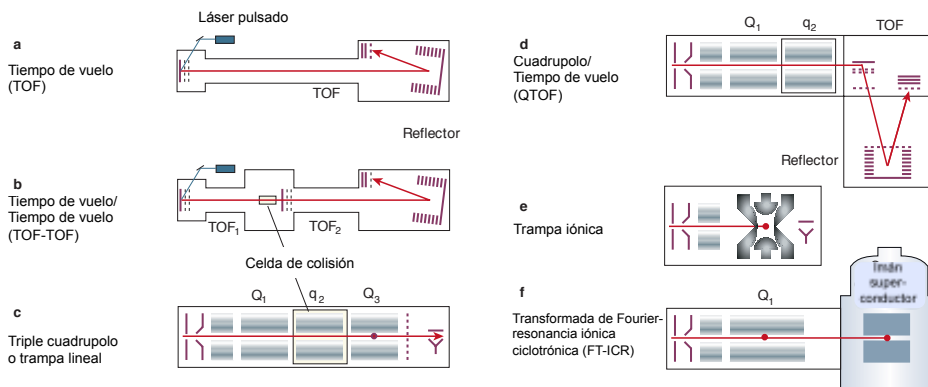


Figura 20. Tipos de analizadores de masas más empleados en proteómica. **a.** Analizador de tiempo de vuelo (TOF), que separa los iones en base a sus diferentes velocidades. **b.** Analizador TOF-TOF, que incorpora una celda de colisión: los iones se seleccionan en el primer TOF, se fragmentan en la celda y se separan en el segundo TOF. **c.** Analizador de triple cuadrupolo, que separa los iones por variaciones en el tiempo de campos eléctricos entre cuatro barras paralelas que actúan de electrodos, los iones se seleccionan en el primer cuadrupolo (Q_1), se fragmentan en el segundo (q_2) y se separan en el tercer cuadrupolo (Q_3); en el analizador de trampa lineal, los iones se capturan en el tercer cuadrupolo (Q_3) y los fragmentos se escanean por resonancia con el campo eléctrico aplicado. **d.** Analizador cuadrupolo/tiempo de vuelo (QTOF), que combina los dos primeros cuadrupolos del triple cuadrupolo con un tiempo de vuelo que mide la masa de los iones. **e.** Analizador de trampa iónica, que captura los iones como en la trampa lineal, fragmenta los iones de determinada m/z y los escanea para dar lugar a su espectro de masas. **f.** El analizador de masas de transformada de Fourier y resonancia iónica ciclotrónica (FT-ICR), atrapa los iones con la ayuda de fuertes campos magnéticos, la figura muestra la combinación de un FT-MS con una trampa lineal. Figura adaptada de Aebersold y Mann [69].

Analizador por tiempo de vuelo (TOF)

El espectrómetro de masas de tiempo de vuelo (TOF) es uno de los más empleados en proteómica por su elevada sensibilidad y resolución. Este analizador mide el *ratio* m/z en base al tiempo que tardan los iones en llegar al detector. Debido a que los iones que entran en el analizador tienen (idealmente) la misma energía cinética, de acuerdo con la ecuación $E_c = (1/2)mv^2$, las velocidades varían inversamente proporcional a sus masas. Por ello, los iones más pesados tienen una velocidad menor y tardan más tiempo en llegar al detector, mientras que los

iones más pequeños tienen una velocidad mayor y son los primeros en llegar al detector. Las pequeñas diferencias en la energía cinética inicial de los iones se compensan añadiendo un reflector (también llamado espejo), que es un campo eléctrico que devuelve los iones con una dirección opuesta a la inicial y a determinado ángulo (ver la figura 20a.), alargando el espacio que tienen que recorrer los iones. Estos analizadores evolucionaron con el tiempo, introduciendo una celda de colisión para producir la fragmentación de los iones que se seleccionan en un primer tiempo de vuelo (TOF_1), y posteriormente separar los fragmentos generados en un segundo tiempo de vuelo (TOF_2), obteniendo los espectros de fragmentación denominados MS/MS. Este último tipo de instrumentos se conocen como TOF-TOF y tienen una mayor resolución y exactitud en las masas. La fragmentación o disociación se produce por la colisión de los iones con un gas inerte que se introduce en la celda de colisión (CID), generalmente nitrógeno. Este tipo de analizador se suele emplear con la fuente de ionización tipo MALDI.

Analizador de triple cuadrupolo

Un cuadrupolo está formado por el conjunto de cuatro barras cilíndricas paralelas que actúan como electrodos. Las barras opuestas se conectan eléctricamente, un par está unido al polo positivo de una fuente variable de corriente continua y el otro par se une al terminal negativo (figura 21). Además, se aplican a cada par de barras potenciales de corriente alterna de radiofrecuencia variables. Una vez que los iones se encuentran entre las barras, las tensiones de corriente continua y alterna se modifican de modo que sólo los iones de determinada relación m/z alcanzan al detector, mientras que el resto de iones inciden en las barras y se convierten en

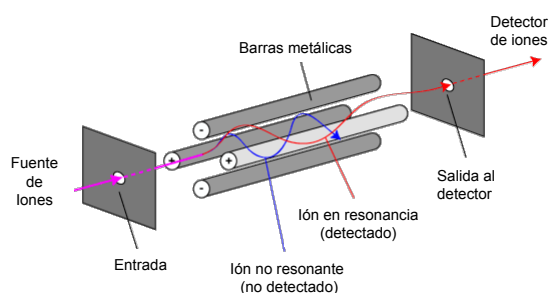


Figura 21. Espectrómetro de masas cuadrupolar.

moléculas neutras. Los dos tipos de espectrómetros de masas de cuadrupolo más comunes son el de cuadrupolo simple y el de triple cuadrupolo (figura 20c). El más empleado en proteómica es el triple cuadrupolo por su capacidad de fragmentar los iones obteniendo los espectros MS/MS. Para ello, el espectrómetro debe trabajar

en dos modos, primero los iones atraviesan los tres cuadrupolos obteniendo las relaciones m/z de todos los iones, y en el segundo modo, el primer cuadrupolo filtra los iones que posteriormente se fragmentan en el segundo cuadrupolo y se escanean en el tercer cuadrupolo.

Analizador de trampa de iones

El analizador de trampa de iones es uno de los analizadores más empleados por su capacidad para aislar y fragmentar iones de péptidos de muestras complejas. En un analizador de trampa de iones los iones gaseosos pueden quedar confinados durante largos períodos de tiempo por la acción de campos eléctricos y/o magnéticos. Para realizar la fragmentación, se seleccionan los iones por medio de voltajes que expulsan el resto de los iones de la trampa, y luego se aumentan los voltajes para incrementar la energía de los iones atrapados, de forma que colisionan con el gas He_2 y se fragmentan. Estos fragmentos se escanean de

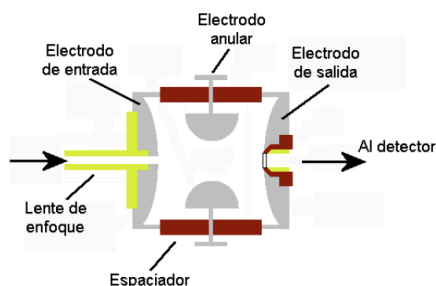


Figura 22. Espectrómetro de masas de trampa de iones.

acuerdo con su relación m/z . Los fragmentos también se pueden volver a atrapar y someter a nuevas fragmentaciones, obteniendo espectros tipo MS/MS/MS o MS^n , aunque este modo no se suele emplear para identificación rutinaria de proteínas, pero sí en identificaciones de péptidos modificados (como péptidos fosforilados). Este tipo de analizador suele acoplarse a una fuente de ionización tipo ESI. En la

figura 22 se representa una trampa de iones sencilla. Un tipo especial de trampa de iones es el orbitrap (trampa de iones orbital), en donde los iones son atrapados bajo un campo eléctrico estático, y orbitan alrededor de un electrodo central y oscilan en dirección axial.

Analizador de transformada de Fourier y resonancia iónica ciclotrónica (FT-ICR)

El espectrómetro de masas FT-ICR funciona de forma similar a una trampa iónica, pero empleando campos magnéticos elevados. Los iones gaseosos se mueven en una trayectoria circular en un plano perpendicular a la dirección del campo

magnético, aplicado con una frecuencia conocida como la frecuencia ciclotrónica. Cuando se aplica un campo eléctrico que oscila cerca de la frecuencia ciclotrónica de los iones atrapados, los iones se excitan alargando su órbita y se miden sus frecuencias con un detector situado en el lado opuesto a la trampa. La energía aplicada se puede ajustar para producir fragmentación o para expulsar a los iones de la trampa. Las frecuencias detectadas se convierten en valores m/z mediante la transformada de Fourier. Estos espectrómetros son los que proporcionan mayor resolución, sensibilidad y exactitud en la masa para medidas de péptidos y proteínas (ver tabla 7).

Analizadores de masas híbridos

En los últimos años, los analizadores se han ido modernizando con el objetivo de incrementar su sensibilidad, su exactitud en las masas y su capacidad de tratar más cantidad de muestras a la vez. Con esta finalidad han surgido los espectrómetros de masas híbridos, que están formados por la combinación de más de un espectrómetro de masas. En la tabla 7 podemos ver las combinaciones más empleadas en proteómica con las características que proporcionan.

Instrumento	Aplicaciones	Resolución	Exactitud	Sensibilidad	Rango dinámico	Rapidez de escaneo
LIT (LTQ)	Identificación	2000	100 ppm	Femtomoles	1×10^4	Rápida
TQ	Cuantificación	2000	100 ppm	Atomoles	1×10^6	Moderada
LTQ-Orbitrap	Identificación, cuantificación y PTM	100 000	2 ppm	Femtomoles	1×10^4	Moderada
LTQ-FTICR Q-FTICR	Identificación, cuantificación y PTM	500 000	< 2 ppm	Femtomoles	1×10^4	Baja
Q-TOF IT-TOF	Identificación y PTM	10 000	2-5 ppm	Atomoles	1×10^6	Moderada Rápida
Q-LIT	Cuantificación	2000	100 ppm	Atomoles	1×10^6	Moderada Rápida
TTOF	Identificación y cuantificación	40 000	< 2 ppm	Atomoles	1×10^5	Rápida

Tabla 7. Características de los espectrómetros de masas híbridos más empleados en proteómica. LIT: trampa de iones lineal; LTQ: trampa lineal cuadrupolar; TQ: triple cuadrupolo; Orbitrap: trampa orbital; FTICR: transformada de Fourier resonancia iónica ciclotrónica; Q: cuadrupolo; TOF: tiempo de vuelo; IT: trampa de iones; TTOF: triple TOF; PTM: modificaciones postraduccionales. Tabla modificada de Yates y colaboradores [101].

1.3.7.3. Identificación de proteínas

La identificación de proteínas mediante espectrometría de masas se realiza a través de sus péptidos, obtenidos generalmente por digestión con un enzima (tripsina). Uno de los métodos más empleados es el que se conoce por el nombre de huella peptídica (peptide mass fingerprint), y está basado en el espectro de los péptidos que se obtienen tras la digestión de la proteína [117]. La identificación de una proteína se realiza comparando las masas obtenidas con las masas resultantes de una digestión *in silico* a partir de una base de datos proteómica o genómica [118-120].

En el caso de muestras complejas, como el suero, la huella peptídica no proporciona una identificación clara, ya que el espectro de masas (MS) tiene demasiados picos de diferentes péptidos. Por ello, para la identificación de proteínas en estas muestras se suele emplear el espectro de fragmentación, que se obtiene seleccionando los péptidos con una relación m/z determinada y aislándolos en una celda para fragmentarlos mediante disociación inducida por colisión (CID). Una vez fragmentado el péptido, se obtienen las masas de los fragmentos en un espectro de masas en tándem (MS/MS), que proporciona información sobre su secuencia de aminoácidos. Y como en el anterior caso, las masas obtenidas se comparan con los espectros de fragmentación *in silico* calculados a partir de la secuencia de aminoácidos de las proteínas en bases de datos.

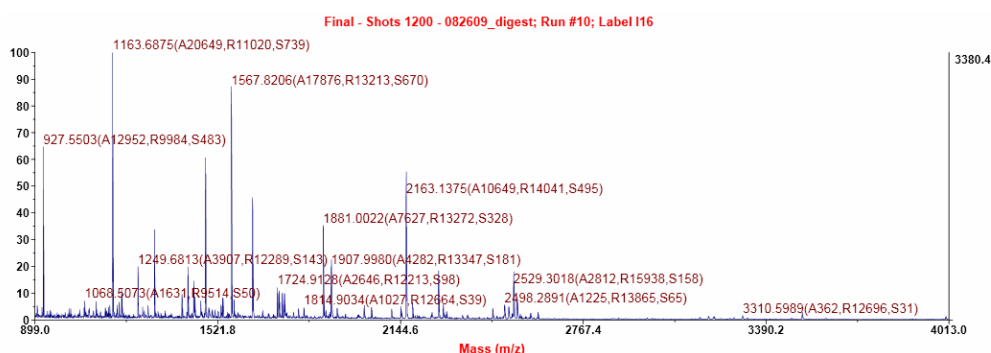


Figura 23. Espectro de masas de la proteína albúmina digerida con tripsina y analizada en un espectrómetro de masas con desorción/ionización con láser asistida por matriz y analizador de tiempo de vuelo (MALDI-TOF).

2. OBJETIVOS

2. Objetivos

Existen trabajos orientados a la búsqueda de marcadores proteicos de artrosis mediante espectrometría de masas, que han aportado una gran cantidad de información. Sin embargo, en la mayoría de ellos no se ha podido avanzar hacia la aplicación clínica de los resultados. Por ello, en este trabajo se ha planteado buscar un método proteómico alternativo con más posibilidades de trasladar a la práctica clínica, y de este modo el principal objetivo de esta tesis ha sido desarrollar un método rápido y económico para la búsqueda de biomarcadores de artrosis y otras enfermedades reumáticas en suero.

Del objetivo principal se derivan los siguientes objetivos:

1. Desarrollar un nuevo método de depleción de los sueros basado en precipitaciones químicas.
2. Aplicar dicho método para comparar el proteoma de sueros de pacientes artrósicos con el proteoma de sueros de donantes sin síntomas.
3. Mejorar la identificación de las proteínas del suero de pacientes con artrosis mediante espectrometría de masas.
4. Crear un método sencillo de separación y análisis de las proteínas del suero, y validar su utilidad para la búsqueda de biomarcadores de enfermedades reumáticas en suero.

3. CAPÍTULO

3. Capítulos

3.1. Artículo I: Comparación de la depleción y la ecualización para reducir las proteínas más abundantes en suero humano

Las proteínas más abundantes del suero o plasma son generalmente un problema en la búsqueda de biomarcadores nuevos o conocidos. Esto se debe a que las proteínas más abundantes representan aproximadamente el 95% de la masa total de proteínas presentes en el plasma. Como consecuencia, se produce un enmascaramiento de las proteínas menos abundantes haciendo muy difícil su detección, y suele ser en esta fracción del suero donde se encuentran los biomarcadores. Para solucionar este problema se han propuesto diferentes estrategias basadas en la reducción o eliminación de dichas proteínas, llamadas depleciones, la más empleada es la cromatografía de inmunoadfinidad. Adicionalmente, se han diseñado otras estrategias para abarcar el rango dinámico de concentración de las proteínas en suero manteniendo el proteoma completo mediante kits comerciales (www.bio-rad.com/proteominer/). El problema de estos métodos es su elevado coste, por lo que recientemente se han publicado alternativas más económicas, sencillas y rápidas mediante precipitaciones con compuestos químicos comunes como el acetonitrilo (ACN) y el ditiotritol (DTT).

El objetivo de este trabajo fue comparar las depleciones químicas empleando ACN y DTT, con un kit comercial diseñado para abarcar el rango dinámico de la concentración de las proteínas en suero mediante librerías de ligandos específicos (ProteoMiner Kit de BIORAD).

Para la correcta cuantificación de las proteínas en nuestras muestras, se hizo un estudio de los diferentes métodos de cuantificación empleados en nuestro laboratorio: el método de Bradford, el kit de BCA (ácido bicinonónico) y la medida de absorbancia a 280nm en el espectrofotómetro NanoDrop UV-Vis. Se hicieron medidas por quintuplicado de una mezcla de 4 proteínas estándar con los tres métodos. Los resultados obtenidos en cuanto a la precisión fueron mejores con el BCA y el NanoDrop, mientras que el Bradford proporcionó la mejor exactitud por lo que se escogió este último método para llevar a cabo el experimento.

La comparación de los métodos de depleción se realizó con la misma mezcla de 5 sueros. Dicha mezcla se trató por cuadruplicado con cada método. Las depleciones con ACN y DTT se hicieron de acuerdo con los protocolos previamente publicados por Kay y colaboradores, y Warder y colaboradores respectivamente, y la equalización se llevó a cabo con el kit ProteoMiner de acuerdo con las instrucciones del fabricante (Bio-Rad). Las proteínas resultantes se separaron por electroforesis monodimensional en geles de acrilamida/bisacrilamida. A continuación, los geles se tiñeron con azul de Coomassie y las bandas de cada carril se digirieron con tripsina acelerada por ultrasonidos. Finalmente, las proteínas correspondientes a cada banda del gel se identificaron mediante espectrometría de masas tipo MALDI-TOF/TOF.

Los resultados obtenidos con la precipitación con ACN mostraron que las proteínas mayores de 72 kDa fueron eliminadas, se obtuvo un 24% de proteínas abundantes de un total de 21 proteínas identificadas y las muestras resultaron enriquecidas en apolipoproteínas. En la precipitación con DTT se identificaron un total de 38 proteínas, de las cuales el 47% pertenecían al grupo de las proteínas abundantes en suero. Las muestras tratadas con este método resultaron enriquecidas en inmunoglobulinas. Y finalmente, con el ProteoMiner kit se identificaron un total de 29 proteínas con un 31% de proteínas abundantes. Los tres métodos presentaron una buena reproducibilidad y sólo se identificaron 4 proteínas comunes a los tres.

Carolina Fernández¹
 Hugo M. Santos²
 Cristina Ruiz-Romero¹
 Francisco J. Blanco¹
 José-Luis Capelo-Martínez^{2,3}

¹Laboratorio de Investigación Osteoarticular y del Envejecimiento, Unidad de Proteómica-Nodo Asociado a ProteoRed-(Genoma España), Centro de Investigación Biomédica, Servicio de Reumatología, INIBIC, Complejo Hospitalario Universitario de A Coruña, Coruña, Spain

²REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, Caparica, Portugal

³BIOSCOPE Group, Physical Chemistry Department, Science Faculty, University of Vigo at Ourense Campus, Ourense, Spain

Received January 1, 2011

Revised May 28, 2011

Accepted June 13, 2011

Research Article

A comparison of depletion versus equalization for reducing high-abundance proteins in human serum

In this work three methods to diminish the content of most highly abundant proteins in human serum have been studied and compared. Protein depletion with ACN or DTT and protein equalization with the ProteoMinerTM (PM) have been assessed by 1-D gel electrophoresis and MS. After treatment 5, 18 and 9 major proteins within the 20 most abundant proteins in serum were identified for the ACN, DTT and PM methods, respectively. The ACN method was efficient for depleting high molecular weight proteins, over 75 kDa, resulting in $10 \pm 4\%$ ($n = 3$) of the total protein content remaining in the depleted serum. In addition, 75% of the proteins belonging to the group of the 20 most abundant proteins were not detected, making this depletion strategy a cheap alternative to expensive commercial tools regularly used for removing high abundance proteins from serum. The ACN extract was found rich in apolipoproteins. The dithiothreitol method promotes the precipitation of proteins rich in disulfide bonds, mainly albumin, with $73 \pm 7\%$ ($n = 3$) of the total protein content remaining in the depleted serum, which was found rich in immunoglobulins. The PM method compresses the dynamic range of the serum proteins, rendering an extract containing $16 \pm 2\%$ ($n = 3$) of the total initial protein content. The extract was found to be rich in both apolipoproteins and immunoglobulins. As a general rule the DTT and PM methods provide a compression of the dynamic range of serum protein concentrations while the ACN method allows an effective depletion of the protein fraction above 72 kDa.

Keywords:

ACN / Biomarker discovery / DTT / Protein depletion / Proteominer

DOI 10.1002/elps.201100183



1 Introduction

It is generally accepted that most abundant proteins (MAPs) in plasma or serum are a problem to deal with when searching for new or known biomarkers. This is due to the fact that $<0.1\%$ of the total number of proteins accounts for c.a. 95% of the bulk mass of proteins present in plasma [1, 2]. These proteins are listed in Table 1 of the Supporting Information. As a result, gel- or MS-based proteomics approaches need to circumvent the large signals produced by the MAPs, as they mask or interfere with the detection of proteins belonging to

the low-abundance protein fraction. It is in this fraction that most biomarkers are usually found [2, 3].

To overcome the above problem, different strategies based on the extraction or on the reduction of the MAPs by depletion or precipitation have been proposed [3]. The aim is to reduce the concentration of MAPs at levels in which they do not interfere with the detection of other low-abundance proteins. Furthermore, some modern approaches have been designed to compress the dynamic range of protein concentrations in complex biological samples, the process being possible while maintaining the entire proteome [4]. (http://wolfson.huji.ac.il/purification/PDF/AlbuminRemoval/BIORAD_ProteoMiner.pdf, accessed 18 October 2010).

At present there is no method considered “unique”, i.e. widely accepted by the proteomics community to avoid high-abundance proteins. The reason for this is that the information retrieved from different depletion methods is not the same since the type and characteristics of the proteins depleted are different. To counter this, some companies offer

Correspondence: Dr. José-Luis Capelo-Martínez, Physical Chemistry Department, University of Vigo, Science Faculty at Ourense Campus, E-32004 Ourense, Spain
E-mail: jlcapelom@uvigo.es
Fax: +34-988-387001

Abbreviations: ALBU, serum albumin, 71317; MAP, most abundant protein; PM, ProteoMiner

Colour Online: See the article online to view Figs. 1–3 in colour.

kits to compress the dynamic range of the proteins or to deplete them in plasma, such as in the case of ProteoMiner (PM) or ProteoPrep™ (http://wolfson.huji.ac.il/purification/PDF/AlbuminRemoval/BIORAD_ProteoMiner.pdf, accessed 18 October 2010, http://wolfson.huji.ac.il/purification/PDF/AlbuminRemoval/SIGMA_ProteoPrepInmunDeplKit.pdf, accessed 18 October 2010) (disclaimer: specific company, product and equipment names are given to provide useful information; their mention (reference) does not imply recommendation or endorsement by the authors). Such kits are expensive and cannot be used extensively in most laboratories. Recently some simple, cheap and fast methods for depletion relying on the use of common chemicals such as ACN and DTT have been proposed as alternatives to such expensive commercial kits [5, 6]. The present academic paper reports on a comparison of these methods with a commercial preparation tool, the PM. This kit has been engineered for compression of the dynamic range of protein concentrations in complex biological samples. The kit is based on the selective adsorption of proteins under capacity-limited binding conditions on a solid phase combinatorial ligand library [7–9].

2 Materials and methods

2.1 Standards and reagents

Protein standards for the protein mixture: BSA (69 kDa), holo-transferrin bovine (76–81 kDa) and β -galactosidase (17 kDa) were purchased from Sigma (Steinheim, Germany); Lysozyme (15 kDa) was from Fluka (Buchs, Switzerland).

Molecular weight marker for gel electrophoresis PageRuler Prestained Protein Ladder was purchased from Fermentas (St. Leon-Rot, Germany).

α -Cyano-4-hydroxycinnamic acid puriss for MALDI-MS from Fluka was used as MALDI matrix. Calibration 1, 4700 Proteomics Analyzer Calibration Mixture from Applied Biosystems was used as mass calibration standard for MALDI-TOF/TOF-MS measurements.

Common reagents were used for gel preparation and protein digestion, as explained in the Supporting Information.

Bradford reagent from Sigma (Steinheim, Germany) and BCA™ Protein Assay Kit from Pierce (Rockford, USA) were used for protein quantification.

ProteoMiner™ Protein Enrichment Small-Capacity Kit was purchased from Bio-Rad (Hercules, CA, USA).

2.2 Serum sample

The human serum samples were obtained from anonymous donors in the Complejo Hospitalario Universitario A Coruña, Spain. Equal volumes of serum from five patients were pooled, aliquoted and stored at -80°C until further

processing. All patients suffered from psoriatic arthritis and they all signed the informed consent. The study was approved by the local ethics committee of Galicia (Spain).

2.3 Apparatus

Gel electrophoresis was performed with an electrophoresis system, model Mini-PROTEAN 3 Cell, from Bio-Rad, following the manufacturer's instructions. The image of the gel after staining was acquired in an ImageScanner from GE Healthcare (Buckinghamshire, UK). Protein digestion was done in safe-lock tubes of 0.5 mL from Eppendorf (Hamburg, Germany). A vacuum concentrator centrifuge model Savant SPD 121 P SpeedVac and a vacuum pump model UVS400A, from Thermo (Waltham, USA) were used for sample drying and sample pre-concentration. An ultrasonic bath was used to speed up the gel washing, the protein reduction and the protein alkylation steps. A sonicator model SONOPULS HD 2200 with a Cup Horn BB6 accessory, from Bandelin (Berlin, Germany) was used to accelerate the enzymatic digestion step. Absorption spectra of serum were recorded as microliter samples using a NANODROP ND-1000 Spectrophotometer from Thermo Scientific (USA).

2.4 Sample treatment

To compare the three quantitation methods studied in this work and described below, mixtures of four standard proteins, BSA, β -galactosidase, holo-transferrin bovine and lysozyme, were prepared by dissolving appropriate amounts of protein standards in water.

2.4.1 Bradford analysis

Mixtures of the four proteins of known concentrations ranging from 0.3 to 2 mg/mL were quantified ($n = 5$) using the Bradford method (modified with HCl) [10]. A BSA standard curve (0–4 mg/mL) was generated in triplicate. Aliquots of 10 μL of each standard and samples ($n = 5$) were transferred to a 96-well microplate, and 90 μL of diluted HCl (8.99 mL water and 9.9 μL HCl 1 M) were added to each well. The plate was shaken for 5 min and then 150 μL of diluted Bradford reagent (2 mL Bradford reagent and 4 mL water) were added. Finally, the plate was shaken again for 5 min before reading at 570 nm in a microplate reader.

2.4.2 BCA analysis

Mixtures of the four proteins of known concentrations ranging from 0.025 to 2 mg/mL were quantified with five replicates using the BCA™ Protein Assay kit [11]. A BSA standard curve (0.02–2 mg/mL) was generated in duplicate. Aliquots of 25 μL of each standard and sample were pipetted into a 96-well microplate and 200 μL of BCA Working

Reagent were added to each well. The plate was shaken on a plate shaker for 30 s. Then the plate was covered and incubated at 37°C for 30 min. Afterwards, the plate was cooled to room temperature, and then the absorbance was measured at 570 nm on a microplate reader.

2.4.3 Nanodrop analysis

Mixtures of the four proteins of known concentrations ranging 1 to 3 mg/mL were quantified ($n = 5$) using a NanoDrop UV-Vis Spectrophotometer [12]. An aliquot of 1 μ L of each sample was pipetted directly onto the pedestal and measured at 280 nm. Calculations of protein concentration from absorbance values were automatically performed by the NanoDrop 2000 software.

2.4.4 Protein depletion with ACN

Protein depletion with ACN was performed according to the protocol described by Kay et al. [5]. A total volume of human serum of 90 μ L, distributed in three aliquots, was depleted with this method. The depletion was done on four independent samples. Three aliquots of 30 μ L were transferred to LoBind microcentrifuge tubes (Eppendorf); each aliquot was diluted with 60 μ L of water and vortexed to mix. ACN (135 μ L) was added to each aliquot and the sample was sonicated for 10 min in an ultrasonic bath. The samples were vortexed briefly and then sonicated for a further 10 min. The protein precipitate was pelleted by centrifugation at $12\,000 \times g$ for 10 min at room temperature. The three supernatants were pooled in a clean LoBind tube and evaporated to dryness in a vacuum concentrator centrifuge without heating. The sample was reconstituted in 50 μ L of buffer with 0.2 M Tris-HCl, 2% w/v SDS, 20% v/v glycerol.

2.4.5 Protein depletion with DTT

Protein depletion with DTT was performed according to the protocol described by Warder et al. [6]. The same initial procedure as described above was performed. Three aliquots of 30 μ L of human serum were depleted in quadruplicate. Fresh DTT, 500 mM (3.33 μ L), was mixed with each aliquot and vortexed briefly. The samples were incubated until a viscous white precipitate persisted (60 min), followed by centrifugation at $14\,000 \times g$ for 20 min. The three supernatants were pooled in a clean LoBind tube and evaporated to dryness in a vacuum concentrator centrifuge. The sample was resuspended in 50 μ L of buffer with 0.2 M Tris-HCl, 2% w/v SDS and 20% v/v glycerol.

2.4.6 Protein depletion with PM

The protein enrichment and depletion of serum samples with the PM kit was carried out according to the manufacturer's instructions (http://wolfson.huji.ac.il/purification/PDF/AlbuminRemoval/BIORAD_ProteoMiner.pdf, accessed 18 October 2010). Human serum (200 μ L) was treated with the

kit in quadruplicate, eluted in 60 μ L of total volume of elution buffer and stored at -20°C until gel electrophoresis.

2.5 Gel electrophoresis

Crude serum samples were quantified with the nanodrop and depleted serum samples were quantified with Bradford assay, due to incompatibility of the elution buffer of the PM kit with the Nanodrop analysis. Amounts of 20 μ g of each one were mixed with 4 μ L of SDS-PAGE loading buffer (10% w/v SDS, Tris-Base 40 mM, pH 6.8, 50% v/v glycerol, 0.1% v/v bromophenol blue, 10% v/v β -mercaptoethanol) in a final volume of 20 μ L. The samples were denatured by heating at 100°C for 5 min and loaded onto 10% polyacrylamide gels with 0.75 mm thickness. Electrophoresis was performed at 120 V (constant voltage) for 65 min. After gel electrophoresis, the gel was stained with Coomassie brilliant blue (CBB) or silver nitrate, according to the following section.

2.6 Stain and image analysis

CBB stain solution was prepared dissolving 1 g of CBB G-250 in 100 mL of a mixture of 50% v/v methanol, 10% v/v acetic acid and 40% v/v water. After gel electrophoresis, gels were rinsed with water and then incubated for 1 h in the staining solution at room temperature. After incubation, the staining solution was removed, and the gel was rinsed with water and then incubated for 1.5 h at room temperature in the destaining solution (40% v/v methanol and 10% v/v glacial acetic acid and 50% water). After destaining, the gel was rinsed with water and incubated for at least 1 h. Images were acquired using an ImageScanner (GE Healthcare).

Silver staining was performed according to standard protocols [13] with modifications. Briefly, gels were fixed in 40% v/v ethanol and 10% v/v acetic acid for 30 min, and washed in water (2×10 min). Sensitization was carried out for 1 min in 0.02% w/v sodium thiosulfate. After two consecutive washes of 1 min each with water, the gels were impregnated with 0.2% w/v silver nitrate in 0.075% w/v formalin for 60–90 min. Excess silver nitrate was removed by washing with water for a few seconds; stain development was carried out in a solution containing 3% w/v sodium carbonate, 12.5 mg/L sodium thiosulfate and 0.025% w/v formalin. Once the bands or spots on the gel reached the desired density, the reaction was stopped by transferring the gels to 3% w/v Trizma-base in 10% v/v acetic acid. After 30 min in the stop solution, gels were scanned and dried or stored at 4°C in water.

2.7 In-gel protein digestion

Ultrasonic in-gel enzymatic digestion was done according to the ultrafast proteolytic digestion protocol previously

developed in our laboratory [14, 15]. Protein bands were manually excised from the gel and placed in safe-lock tubes of 0.5 mL. Protein bands were washed, first with water (100 μ L) and then with ACN (100 μ L), in an ultrasonic bath operating at 35 kHz (100% amplitude) for 5 min for each step. Then, a piece of gel was dried in a vacuum concentrator centrifuge for 5 min. Protein reduction and alkylation steps were included in the protocol to facilitate the enzymatic action and to increase the protein sequence coverage (%). To do so, protein cystine residues were reduced with DTT in an ultrasonic bath operating at 35 kHz (100% amplitude) for 5 min at room temperature. The resulting cysteines were then blocked with iodoacetamide in an ultrasonic bath operating at 35 kHz (100% amplitude) for 5 min at room temperature. After reduction and alkylation steps, the gel was submitted again to the same washing procedure as described above, followed by another drying step of 5 min. Afterwards, trypsin (375 ng in 25 μ L) was added to the dried protein bands and the in-gel protein digestion was performed in the sonicator operating at 50% amplitude for 2 min. Next, trypsin activity was stopped by the addition of 20 μ L of 5% v/v formic acid. The supernatants were collected into new Eppendorfs and to the remaining gel piece, 25 μ L of extraction solution (50% v/v ACN, 0.1% v/v trifluoroacetic acid) were added to each tube and the samples were sonicated for 2 min at 50% amplitude in the sonicator. The new supernatants were pooled with the prior supernatants and were evaporated to dryness and reconstituted with 5 μ L of 0.3% v/v formic acid.

2.8 MALDI-TOF-MS/MS analysis

Prior to analysis, 1 μ L of sample was hand-spotted onto a MALDI target plate (384-spot Teflon[®]-coated plate) and allowed to air dry at room temperature. Subsequently, 1 μ L of a 3 mg/mL solution of α -cyano-4-hydroxycinnamic acid matrix in 0.1% v/v TFA and 50% v/v ACN was added to the dried peptide digest spots and again allowed to air dry. The samples were analyzed using a MALDI-TOF/TOF mass spectrometer model 4800 Proteomics Analyzer from Applied Biosystems (Framingham, MA) equipped with a Nd:YAG laser (λ = 355 nm) and 4000 Series Explorer[™] software (Applied Biosystems). Internal calibration, data acquisition, processing and interpretation were carried out as recommended by the manufacturer. All mass spectra were externally calibrated using a standard peptide mixture (Applied Biosystems).

Mass spectral analysis for each sample was based on an average of 1500 laser shots. Fragmentation spectra was acquired by selecting the 20 most-abundant ions of each MALDI-TOF peptide mass map and averaging 2000 laser shots per fragmentation spectrum. For MS/MS all peaks with an $S/N > 10$ were included in the database search. The analysis of MS/MS spectra was performed by applying the MASCOT (Matrix Science) search engine with the following parameters: (i) Swiss-Prot, 2010 Database (516 081 sequences; 181 677 051 residues); (ii) taxonomy, *Homo sapiens*

(20 292 sequences); (iii) enzyme, trypsin; (iv) one missed cleavage; (v) fixed modifications, carbamidomethyl (C); (vi) variable modifications, oxidation (M); (vii) peptide mass tolerance up to 50 ppm; (viii) fragment mass tolerance up to 0.3 Da. The criteria used to accept the identification was significant homology scores achieved in MASCOT ($p < 0.05$).

3 Results and discussion

The workflow of the research described below is depicted in Supporting Information Fig. 1.

3.1 Comparison of methods for the total protein quantification

First, a study was carried out to select the best way to quantify the total protein content in the samples before and after the protein depletion. Second, it was continued by investigating the performance of the three methods of depletion/equalization mentioned above.

Without any further details our results showed that the (i) BCA and nanodrop methods provided similar results in terms of accuracy and precision and that the (ii) Bradford method provided the best results in terms of accuracy, at least under the conditions employed in this study. With these results in our hands, we decided to use the Bradford method to quantify the depleted samples and the nanodrop technique to quantify the crude serum.

A detailed explanation of the set of experiments that were done to compare BCA, Bradford and nanodrop methods can be found in the Supporting Information.

3.2 Protein depletion with ACN

The identification of proteins corresponding to bands of non-depleted serum is shown in Supporting Information Fig. 2. Also, as seen in Supporting Information Fig. 2, high-abundant proteins are identified in almost all bands, most of which are albumin.

Kay et al. [5] described a simple and efficient methodology to deplete proteins by making the plasma samples up to 60% v/v with ACN. To assess the effectiveness of this procedure, we developed a set of experiments to find out the best concentration of ACN to promote protein depletion. The first problem we encountered and overcame was one regarded to the fact that once the sample was pelleted, pellet reconstitution was incomplete (naked eye) in water/TFA (0.1% v/v). Therefore, next we tried a solution of 0.2 M Tris-HCl at pH 6.8 prepared in 2% w/v SDS and 20% v/v glycerol. This solution resulted in a better reconstitution of the pellet. Consequently, this was the solution we used to reconstitute the pellets in further experiments. Afterwards, a series of samples of 20 μ L of the same serum were prepared according to the protocol described in Section 2.4.4. Serum

samples were made up with ACN so its concentration was varied v/v between 47 and 60% as shown in Fig. 1.

In agreement with Kay et al., we found that the ACN concentration is best kept at 57% v/v; otherwise the depletion of the protein fraction above 72 kDa is not efficiently achieved, as demonstrated by the visible bands in the 1-D gel (Fig. 1). It was also found that this is a robust method, because of its high reproducibility, as shown in Supporting Information Fig. 3.

In this figure different gels show similar densities of CBB-stained protein bands for the ACN method. These results were later confirmed by MS (see further sections below). Protein precipitation has been suggested to occur because of the addition of ACN to an aqueous solvent, which weakens the hydrophobic interaction and enhances the peptide–peptide hydrogen bond leading to the denaturation of proteins. Once denatured the proteins tend to aggregate and precipitate [16].

A total of (see Supporting Information Table 2) 21 proteins were identified. Of these, five belonged to the group of the 20 MAPs, i.e. serum albumin, 71317 (ALBU), APOA1 (apolipoprotein A-I, 30759) and APOA2 (apolipoprotein A-II, 11282), A1AT (α -1-antitrypsin, 46878) and A1AG1 (α -1-acid glycoprotein 1, 23725). This means that the concentration of the other 15 proteins in the top 20 was reduced to levels below our limit of detection. In this way the relative protein in-gel load of other less-abundant proteins was increased. Six of the 21 proteins identified were apolipoproteins. This result

is in agreement with the one reported by Kay et al., and it further supports their hypothesis that this method selectively enriches proteins of a hydrophobic nature. The apolipoproteins remain in solution probably because of their amphipathic properties. Apolipoproteins have been linked to cardiovascular-related diseases and are seen as risk markers of myocardial infarction [17]. Therefore, protein precipitation with ACN could be recommended as a potential cheap and fast method to avoid high-abundance proteins in studies related to cardiovascular diseases.

The gel band corresponding to albumin was less dense in the depleted serum than in the crude sample (see Fig. 2). This suggests that the majority of this protein was depleted.

Regarding reproducibility, as shown in Fig. 3, the same nine proteins were identified (three gels) in four replicates, six proteins in three replicates; three proteins in two replicates and three proteins in one replicate. The amount of protein remaining in the serum supernatant was found to be $10 \pm 4\%$ ($n = 3$) of the initial total protein content in the crude serum. The number of bands excised including the proteins identified in each band is depicted in Supporting Information Fig. 4.

3.3 Protein depletion with DTT

Recently, Warder et al. reported a robust, fast and easy method to deplete high-abundance disulfide-rich proteins in

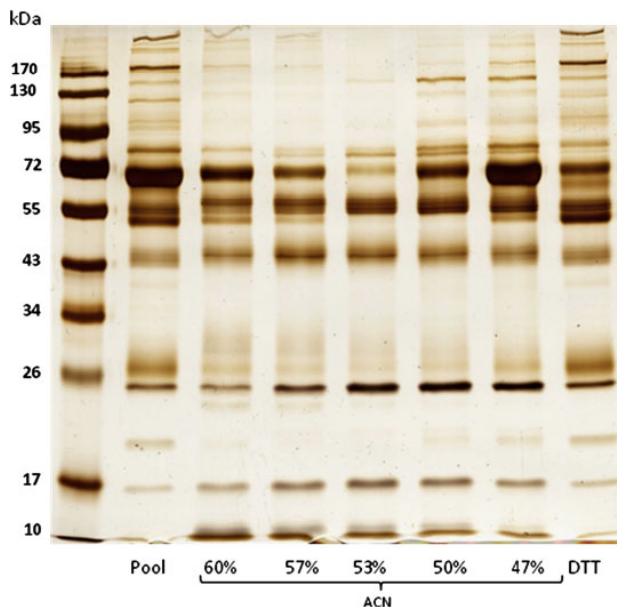


Figure 1. Silver stained 1-D PAGE of serum samples after (i) protein depletion with different ACN concentrations, (ii) supernatant evaporation to dryness in a vacuum concentrator centrifuge without heating and (iii) reconstitution in a solution of 0.2 M Tris-HCl, 2% w/v SDS and 20% v/v glycerol.

human plasma [6]. We used this methodology with minor modifications. While these authors analyzed the extract of the sample after depletion with DTT, we pooled the extracts of three depleted sera samples; then the samples were evaporated to dryness in a vacuum concentrator centrifuge, and finally the sample was reconstituted in 50 μ L of buffer with 0.2 M Tris-HCl, 2% w/v SDS and 20% v/v glycerol. The sample obtained after reconstitution was run in a 1-D SDS PAGE. The result is shown in Fig. 2, where it can be clearly seen that the depletion is reproducible for both the number and density of bands. The gel profile is in accordance with the one presented by Warder et al. [6]. It was also found that this method reduces the presence of ALBU, relative to the crude sample. DTT is commonly used to disrupt intramolecular and intermolecular disulfide bonds, helping to unfold proteins rich in disulfide bonds. Such unfolded proteins tend to form aggregates, which finally precipitate. It is suggested that this was the mechanism with which albumin and serotransferrin, both high-abundance disulfide-rich proteins, were preferentially depleted. The number of proteins identified in the depleted serum was 38, of which 18 belonged to the group of the 20 MAPs [6]. Fourteen of the proteins identified out of the top 20 group were immunoglobulins. These proteins are mainly involved in the secondary immune response and their related patterns obtained from different physiological or pathological conditions are used for research and diagnostic purposes.

Therefore, this method of depletion could potentially be used to aid the study of diseases that are characterized by modifications in the content of immunoglobulins. Such diseases include multiple myeloma, discrasias, Waldenström's macroglobulinemia, amyloidosis, solitary plasmacytoma, heavy-chain diseases, monoclonal gammopathy of undetermined significance and antibody deficiency disorders [18].

Warder et al. found the gel bands above 95 kDa highly rich in ALBU [6]. However, in our case a different set of proteins was found, as follows: APOB (apolipoprotein B-100, 516666), CO4-A, CO3 (complement C3, 188569), A2MG (α -2-macroglobulin, 164614), CERU (ceruloplasmin, 122983) and ITIH2 (inter- α -trypsin inhibitor heavy chain H2, 106853).

It is hypothesized in this paper that this difference might be related to the modification that was done to the protocol, as described in the beginning of this section. Warder et al. [6] have suggested that the protein depletion promoted by DTT can be directly linked to the protein's cysteine content. This proposal seems logical as the depleted proteins TRFE (Serotransferrin, 79280) and ALBU have a high ratio total cysteines/total amino acids of 4.9 and 5.4%, respectively. However, it is noteworthy that some non-depleted proteins identified in bands 9 and 14 (Supporting Information Fig. 5) have similar total cysteines/total amino acids ratios. This is the case, for instance, for the following

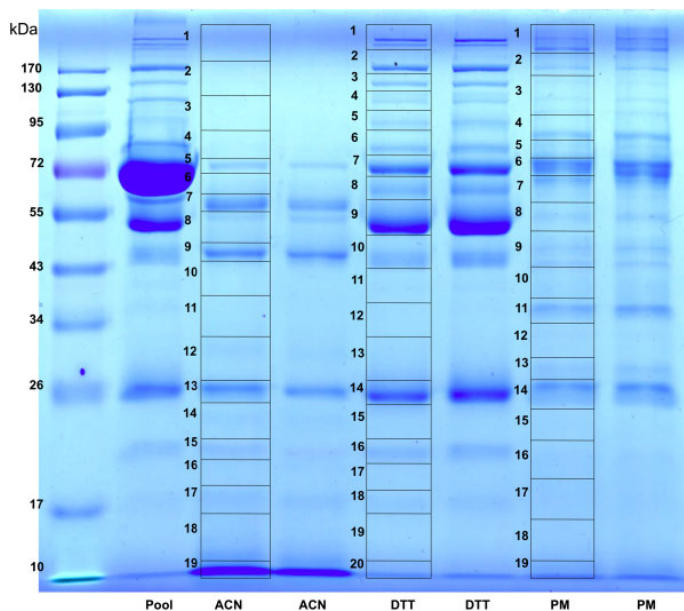


Figure 2. CBB-stained 1-D PAGE of serum samples after protein depletion with (i) 57% v/v ACN, (ii) 500 mM DTT and (iii) using the PM. For further details refer to sample treatment section.

proteins: IGHG2 (Igγ-2 chain C region, 36505), 3.4%; FETUA (α-2-HS-glycoprotein, 40098), 3.5% and IGHG3 (Igγ-3 chain C region, 42287), 4.5%.

To explain the difference between one efficient process of depletion and another that is equally efficient, for the proteins mentioned above, it is suggested in this paper that their relative abundance could be another factor to take into consideration. This can happen due to the fact that albumin and transferrin (two proteins in ca. 10 000 sera proteins) constitute 54.3 and 3.3% of the total serum protein content.

The amount of protein remaining in the serum supernatant after the depletion was found around $73 \pm 7\%$ ($n = 3$). Figure 3 shows that, of the 38 identified proteins, 80% were found in at least two of the four 1-D PAGE replicates performed. The number of bands excised, including the proteins identified in each band, is depicted in Supporting Information Fig. 5.

3.4 Protein equalization with PM

PM is the name given to a methodology specifically developed by BioRad, which is used for the compression of the dynamic range of protein concentrations in complex biological samples such as serum. The company claims that the use of this technology allows detection of medium- and low-abundance proteins. The basis of the PM is a large, highly diverse library of hexapeptides bound to chromatographic supports that are mixed with the complex protein samples. Because the bead capacity limits binding capacity, high-abundance proteins quickly saturate their ligands and excess protein is washed out during the procedure. In contrast, low-abundance proteins are concentrated on their specific ligands. The result is the decreasing of the dynamic range of proteins in the sample. This fact can be observed in Fig. 2, where the pattern of the gel bands corresponding to

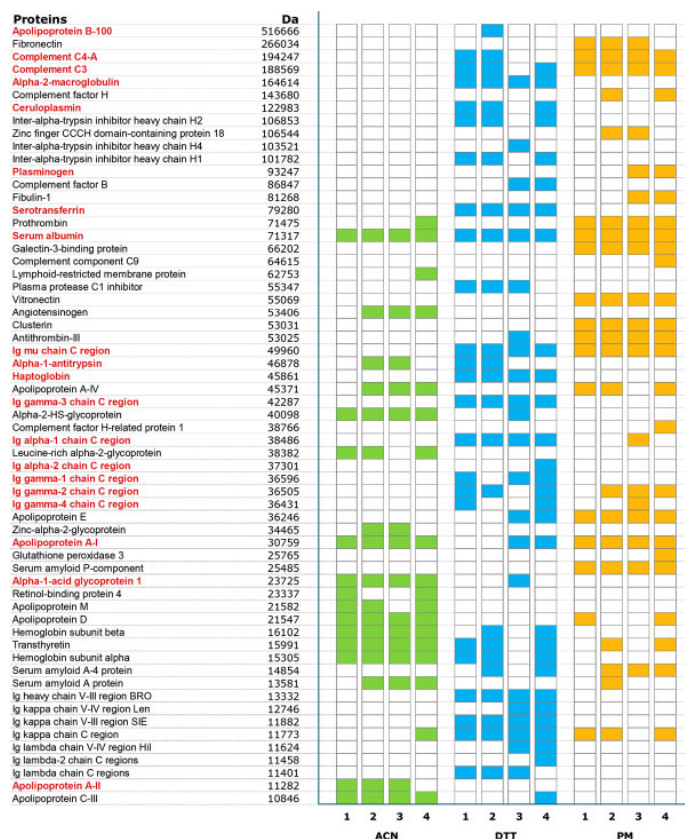


Figure 3. Proteins identified for each method of depletion. Each column corresponds to a gel line. Lines 1 and 2 were run in gel number 1, lines 3 and 4 were run in gels 2 and 3, respectively.

the PM is almost the same as the one present in the bands corresponding to the crude sample, but with less density. The obtained protein pattern was very reproducible as it was the same for the four replicates running in three gels, which were prepared in different days.

The protein distribution in the gel clearly shows that the MAPs were not totally depleted, although it may be seen, as well, that the content in ALBU (72 kDa) and proteins present in the 55 kDa (band 8) and 26 kDa (band 14) have been drastically decreased, compared with the crude sample. With the help of the PM method, 29 proteins were identified, among which nine belonged to the 20 MAPs. The pattern of the proteins identified (see Fig. 3) comprised immunoglobulins and apolipoproteins, thus rendering a part of the serum proteome that was not reached using the ACN or the DTT methods alone. The amount of protein remaining in the serum supernatant was found to be $16 \pm 2\%$ ($n = 3$) of the initial total protein content in the crude serum.

Figure 3 shows that, of the 29 proteins identified, 12 were identified in four replicates, five in three replicates, six in two replicates and six in one replicate. In Supporting Information Fig. 6, the number of bands excised including the identified proteins in each band is depicted. These results confirm the robustness of the method.

3.5 Comparison of the three procedures

ACN was the method that rendered the extract with the lowest protein content, $10 \pm 4\%$ ($n = 3$), while the use of DTT resulted in an extract with a protein content of $73 \pm 7\%$ ($n = 3$). The PM equalizes the serum proteins rendering an extract with a total protein content of $16 \pm 2\%$ ($n = 3$). Five, nine and 18 proteins included in the group of the most-abundant ones (top 20) were identified after depletion with ACN, PM and DTT methods, respectively. This suggests that efficient depletion with ACN removes such proteins from the serum. On the other hand, the DTT method seems to compress the dynamic range of protein concentration in serum in such a way that almost all proteins on the top 20 are detected. And yet the number of other proteins identified is as high as the one obtained with the PM or ACN.

Band 6 for PM and band 7 for DTT in Fig. 2 can be used as examples of the differences in the compression of the dynamic range of the serum proteins concentration. In the case of the DTT method, the protein ALBU was identified, but COA4 was also identified for the PM method. The reasons for this to happen are the better ALBU depletion achieved through the PM method, which makes it possible to identify both proteins in the same MALDI spectrum.

The ACN-depleted serum was found rich in apolipoproteins; so when DTT was used, the extract was found abundant in immunoglobulins. The PM depletion rendered an extract rich in both immunoglobulins and apolipoproteins.

Supporting Information Figure 7 shows the Venn diagram of the proteins identified with the methods that we have compared in this work.

It must be emphasized that only four proteins were identified within the three procedures, thus demonstrating that the part of the serum proteome that can be investigated using each method is not the same.

It is interesting to note that each method provides different information. The ACN method favored the precipitation of the proteins with molecular weights >72 kDa while the PM method seems to be better at depleting proteins with molecular weights <26 kDa. As a general trend, the DTT and the PM methods allow the dynamic range of the MAPs in serum to be compressed while maintaining a reasonable range of other less-abundant proteins.

4 Concluding remarks

Three methods utilized to remove major proteins or decrease their relative abundance in human serum have been compared. When the ACN method was used, only five proteins of the 20 most-abundant serum proteins were identified, thus showing the method's effectiveness of the depletion of major proteins. The DTT and PM methods did not remove major proteins as thoroughly as the ACN method did, 18 and nine proteins out of the 20 most-abundant serum proteins were identified, respectively. However, DTT and PM seemed to work better to compress the dynamic range of the protein concentrations. Excluding the proteins belonging to the 20 MAPs, with the help of both methods, 20 more proteins were identified while with the ACN method, only 16 were identified. It was found that serum depleted with ACN was rich in apolipoproteins and, therefore, this method is proposed as a cheap strategy to use in studies dealing with depletion in cardiovascular-related diseases. Serum depleted with DTT was found rich in immunoglobulins and for this reason, it is proposed as an economic method in studies of diseases that are characterized by modifications of such proteins, like multiple myeloma.

This work highlights the importance of understanding that different parts of the serum/plasma proteome are investigated as a function of the strategy chosen to reduce/remove high abundant proteins. Therefore, it is recommended that different strategies of reduction/depletion be used in the same study in order to guarantee that the serum proteome is investigated to the largest extent possible.

H. M. S acknowledges the post-doctoral grant SRFH/BPD/73997/2010 provided by FCT (Science and Technological Foundation). C. F. acknowledges the Foundation of the Complexo Universitario de A Coruña. Dr. J.-L. C.-M. is grateful to the Xunta de Galicia (Spain) for the program Isidro Parga Pondal and for financial support provided under project 09CSA043383PR.

The authors have declared no conflict of interest.

5 References

- [1] Stéphane, R., Laurent, T., Monique, P., Marie-Thérèse, P., Sylvain, L., *Proteome Sci.* 2006, 4, 1–7.
- [2] Anderson, N. L., Anderson, N. G., *Mol. Cell. Proteomics* 2002, 1, 845–867.
- [3] Stéphane, R., Laurent, T., Monique, P., Martial, S., Marie-Thérèse, P., Patrick, J., Sylvain, L., *J. Proteomics* 2009, 72, 945–951.
- [4] Righetti, P. G., Castagna, A., Antonioli, P., Boschetti, E., *Electrophoresis* 2005, 26, 297–319.
- [5] Kay, R., Barton, C., Ratcliffe, L., Matharoo-Ball, B., Brown, P., Roberts, J., Teale, P., Creaser, C., *Rapid Commun. Mass Spectrom.* 2008, 22, 3255–3260.
- [6] Warder, S. E., Tucker, L. A., Strelitzer, T. J., McKeegan, E. M., Meuth, J. L., Jung, P. M., Saraf, A., Singh, B., Lai-Zhang, J., Gagne, G., Rogers, J. C., *Anal. Biochem.* 2009, 387, 184–193.
- [7] Guerrier, L., Thulasiraman, V., Castagna, A., Fortis, F., Lin, S., Lomas, L., Righetti, P. G., Boschetti, E., *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, 2006, 833, 33–40.
- [8] Simó, C., Bachi, A., Cattaneo, A., Guerrier, L., Fortis, F., Boschetti, E., Podtelejnikov, A., Righetti, P. G., *Anal. Chem.* 2008, 80, 3547–3556.
- [9] Simó, C., Bachi, A., Cattaneo, A., Guerrier, L., Fortis, F., Boschetti, E., Masseroli, M., Righetti, P. G., *Anal. Chem.* 2008, 80, 3557–3565.
- [10] Kruger, N. J., *The Bradford Method for Protein Quantitation. The Protein Protocols Handbook*, 2nd Ed., Humana Press, New York, USA 2002.
- [11] Walker, J. M., *The Bicinchoninic Acid (BCA) Assay for Protein Quantitation*, 2nd Ed., Humana Press, New York, USA 2002.
- [12] Aitken, A., Learmonth, M., *Protein Determination by UV Absorption. The Protein Protocols Handbook*, 1st Ed., Humana Press, New York, USA 1996.
- [13] Rabilloud, T., Brodard, V., Peltre, G., Righetti, P. G., Ettori, C., *Electrophoresis* 1992, 13, 264–266.
- [14] Rial-Otero, R., Carreira, R. J., Cordeiro, F. M., Moro, A. J., Fernandes, L., Moura, I., Capelo, J. L., *J. Proteome Res.* 2007, 6, 909–912.
- [15] Galesio, M., Vieira, D. V., Rial-Otero, R., Lodeiro, C., Moura, I., Capelo, J. L., *J. Proteome Res.* 2008, 7, 2097–2106.
- [16] Gecko, K., Ohmae, E., Hameyama, K., Takagi, T., *Biochim. Biophys. Acta* 1998, 1387, 195–205.
- [17] Holme, I., Aastveit, A. H., Hammar, N., Jungner, I., Walldius, G., *Atherosclerosis* 2010, 213, 299–305.
- [18] Miller, I., Goldfarb, M., Immunoglobulin patterns in health and disease, in: Smejkal, G. B., Lazarev, A., (Eds.), *Separation Methods in Proteomics*, Taylor and Francis, London, UK 2006.

3.2. Artículo II: Depleción secuencial de suero humano en la búsqueda de biomarcadores de artrosis

La búsqueda de biomarcadores proteicos en proteómica clínica se realiza generalmente mediante técnicas de cromatografía líquida o electroforesis en gel acopladas a espectrometría de masas. Las muestras más utilizadas para la búsqueda de biomarcadores son el suero y el plasma debido a su fácil disponibilidad. Estas muestras tienen más de 10.000 proteínas diferentes aunque sólo unas 20 proteínas representan más del 95% de la concentración total. La señal de estas proteínas abundantes enmascara la señal de otras proteínas menos abundantes que suelen contener información relacionada con diferentes enfermedades. La solución a este problema es eliminar o reducir las proteínas abundantes. Sin embargo, no hay un método universal para solucionar este problema. Existen numerosos métodos en la literatura para deplecionar las proteínas abundantes, como las columnas cromatográficas de inmutioafinidad para HPLC (cromatografía de líquidos de alta eficacia) o diversos kits comerciales. Pero recientemente se están empleando reactivos químicos con este fin, que resultan más económicos. En nuestro trabajo previo comparamos dos precipitaciones químicas con DTT (ditiotreitól) y ACN (acetónitrilo) con un kit comercial para abarcar el rango dinámico de las proteínas en el suero. Una vez deplecionado el suero, la búsqueda de biomarcadores se realiza mediante diversas metodologías, entre las que se encuentra la electroforesis bidimensional. La electroforesis diferencial en gel (2D-DIGE) permite analizar muestras complejas marcadas con diferentes tintes fluorescentes en un mismo gel, evitando el problema de la variabilidad entre geles y permitiendo normalizar las intensidades de las manchas con un único estándar interno.

El objetivo de este trabajo fue desarrollar un nuevo método de depleción basado en la combinación de dos precipitaciones químicas para la búsqueda de biomarcadores de artrosis en suero mediante 2D-DIGE.

Se realizaron 2 conjuntos (pooles), de 10 muestras cada uno, de donantes no sintomáticos control y otros 2 conjuntos de 10 muestras cada uno de pacientes artrósicos (OA) de grado IV. Los sueros se deplecionaron con un nuevo método de depleción secuencial que consta de dos precipitaciones: depleción con DTT

seguida de una depleción con ACN. Las proteínas resultantes se cuantificaron mediante el método de Bradford, y se marcaron con los fluoróforos Cy3 y Cy5 iguales cantidades de proteína para cada condición, para evitar artefactos debidos al marcaje de un tinte específico a una condición, se realizó un marcaje inverso. Una mezcla de las dos condiciones se marcó con el Cy2 como estándar interno. Los cuatro pares Cy3 y Cy5 se combinaron y mezclaron con el estándar Cy2. La primera separación, el isoelectroenfoque, se realizó en tiras con gradiente de pH inmovilizado (IPG) de 7cm y rango de pH: 4-7 no lineal. La segunda separación se realizó en geles del 12.5% de poliacrilamida corridos a la vez en la misma cubeta de electroforesis. Los geles se escanearon en el escáner DIGE Imager obteniendo las imágenes para cada fluoróforo. Finalmente, se realizó el análisis semi-automático de las imágenes con el software Progenesis SameSpots obteniendo una tabla con las manchas moduladas y sus p-valores.

Para saber a qué proteínas pertenecían las manchas moduladas, los geles se tiñeron con Coomassie coloidal, las manchas se cortaron y se digirieron con tripsina acelerada por ultrasonidos. Los péptidos obtenidos se analizaron con un espectrómetro de masas tipo MALDI-TOF/TOF y la identificación de las proteínas se realizó con el software Mascot. Por último, se realizaron inmunodetecciones (western blot) para validar las proteínas moduladas.

Se detectó una media de 210 manchas en los geles, de las cuales 46 estaban significativamente alteradas entre muestras OA y control (29 aumentadas y 17 disminuidas). Las 46 manchas se identificaron mediante espectrometría de masas, y correspondían a 16 formas proteicas. Estas proteínas se encuentran involucradas en diferentes funciones como el transporte de lípidos, respuesta inmune, etc. Curiosamente, tres manchas correspondientes a la misma proteína, haptoglobina (HPT), se encontraron moduladas de diferente forma, dos aumentadas y una disminuida en artrosis frente a controles. Finalmente se confirmó el aumento de la cadena beta de la haptoglobina en pacientes con OA frente a donantes control, mediante análisis por western blot sobre 15 muestras de controles y 15 muestras de artrósicos.

RESEARCH

Open Access

Sequential depletion of human serum for the search of osteoarthritis biomarkers

Carolina Fernández-Costa¹, Valentina Calamia¹, Patricia Fernández-Puente¹, José-Luis Capelo-Martínez^{2,3}, Cristina Ruiz-Romero^{1,4*} and Francisco J Blanco^{1,4*}

Abstract

Background: The field of biomarker discovery, development and application has been the subject of intense interest and activity, especially with the recent emergence of new technologies, such as proteomics-based approaches. In proteomics, search for biomarkers in biological fluids such as human serum is a challenging issue, mainly due to the high dynamic range of proteins present in these types of samples. Methods for reducing the content of most highly abundant proteins have been developed, including immunodepletion or protein equalization. In this work, we report for the first time the combination of a chemical sequential depletion method based in two protein precipitations with acetonitrile and DTT, with a subsequent two-dimensional difference in-gel electrophoresis (2D-DIGE) analysis for the search of osteoarthritis (OA) biomarkers in human serum. The depletion method proposed is non-expensive, of easy implementation and allows fast sample throughput.

Results: Following this workflow, we have compared sample pools of human serum obtained from 20 OA patients and 20 healthy controls. The DIGE study led to the identification of 16 protein forms (corresponding to 14 different proteins) that were significantly ($p < 0.05$) altered in OA when compared to controls (8 increased and 7 decreased). Immunoblot analyses confirmed for the first time the increase of an isoform of Haptoglobin beta chain (HPT) in sera from OA patients.

Conclusions: Altogether, these data demonstrate the utility of the proposed chemical sequential depletion for the analysis of sera in protein biomarker discovery approaches, exhibit the usefulness of quantitative 2D gel-based strategies for the characterization of disease-specific patterns of protein modifications, and also provide a list of OA biomarker candidates for validation.

Background

In clinical proteomics, searching for protein biomarkers is currently done using mass spectrometry-based tools coupled to chromatography or to gel electrophoresis techniques. The samples most commonly used for biomarker discovery are plasma and serum, due to their easy availability. It is generally agreed that each of these samples have more than ten thousand different proteins. From such large amount, it is estimated that circa 20 proteins account for more than 95% of the bulk mass of all proteins [1,2]. These large abundant proteins hinder the searching for putative disease biomarkers, as the

signals belonging to such proteins mask those from other less abundant, yet more important proteins in terms of information related to diseases. Therefore, removal of the most abundant proteins helps to overcome the aforementioned problem. However, to date there is no universal method to accomplish this task. Therefore, many different protocols to deplete high abundant proteins can be found in literature comprising, but not limited to: (i) HPLC columns containing antibodies to the most abundant proteins, (ii) spin columns, (iii) the use of chemical reagents, (iv) or even sophisticated kits used to compress the dynamic range of the proteins rather than to deplete them [3-5].

Recently, the use of non-expensive chemical reagents as a way to deplete high abundance proteins has attracted the attention of the proteomics community [6,7]. Furthermore, it has been suggested to link the use

* Correspondence: cristina.ruiz.romero@sergas.es; fblagar@sergas.es

¹Rheumatology Division, ProteoRed/ISCIII Proteomics Group, INIBIC – Hospital Universitario de A Coruña, 15006 A Coruña, Spain

⁴CIBER-BBN Instituto de Salud Carlos III, A Coruña, Spain

Full list of author information is available at the end of the article



of such methods to the resulting proteome after depletion. For instance, serum depleted with DTT was found rich in IgG-type proteins and for this reason it has been proposed as an economic method in studies of diseases that are characterized by modifications in the IgG-type protein content, such as multiple myeloma [8]. After depletion, the resulting proteome is investigated for biomarker discovery using different methodologies, including gel electrophoresis. Two-dimensional difference in gel electrophoresis, 2D-DIGE, has emerged as a powerful tool in proteomics. In 2D-DIGE, complex protein samples are labeled with different fluorescent dyes, mixed together and separated in the same gel. This method solves the problem of gel-to-gel variation [9], and allows for high sample throughput as only one gel is needed to perform the comparisons and spot intensities can be normalized against a unique internal standard.

The present work reports for the first time the combination of a chemical sequential depletion method combined with 2D-DIGE for the search of osteoarthritis biomarkers. Osteoarthritis is a degenerative joint disease characterized by cartilage destruction and bone changes, occasionally accompanied by synovial inflammation [10]. It is the most common rheumatic pathology, and estimates are that the number of people with some degree of OA will double over the next three decades [11]. Despite this high prevalence, existing diagnostic tests are very limited and ultimately rely on the subjective description of pain symptoms, stiffness in the affected joints, and radiography [12]. These limitations have increased the interest in identifying new specific biological markers for cartilage degradation, both to facilitate early diagnosis of joint destruction and to improve the prognosis and evaluation of disease progression. Besides the promising results obtained with some cartilage-specific proteins, still useful biomarkers are needed for clinical practice, which could be effective for the development of early diagnostic tools and/or alternative therapies [13]. The method proposed of sequential depletion for biomarker search in serum samples is non-expensive, of easy implementation and allows fast sample throughput.

Results and discussion

This work illustrates the application of a chemical sequential depletion method in quantitative proteomics studies for biomarker discovery. The search of novel biomarkers is an area in which proteomics has recently emerged as a powerful approach for the detection of new proteins with diagnostic, prognostic or therapy evaluation utilities. In this field, a number of proteomics experiments have been followed in the last years with the aim of characterizing novel molecules with biomarker power for joint diseases, including osteoarthritis (OA) [14]. In the present study, we report the

combination of a chemical sequential depletion method combined with 2D-DIGE for the search of OA biomarkers. This method reduces the dynamic range of serum proteins, thus allowing the identification of a higher number of proteins in the samples.

Sequential depletion of human serum for proteomics analyses

We have recently reported the comparison of two different chemical depletion methods, using acetonitrile or DTT, with the protein equalization obtained with the ProteoMiner kit (Bio-Rad) [8] for the mass spectrometry analysis of human serum proteins. Our results showed how ACN depletion was efficient for depleting high molecular weight proteins (over 75 kDa), whereas DTT depletion primarily promotes the precipitation of proteins rich in disulphide bonds (mainly albumin). In the present work, we combine these two chemical depletions in a sequential way, with the aim of simplifying the serum protein profile while removing high abundant (usually uninformative) proteins. The workflow followed for this sequential depletion is showed in Figure 1, and is fully described in the Methods section. Human sera were first depleted with DTT, and the supernatant was further treated with ACN in order to efficiently reduce the dynamic range of proteins in our samples. The protein recovery employing this sequential strategy was $2 \pm 0.3\%$, while the use of either DTT or ACN depletion alone leads to a yield of $32 \pm 2\%$ or $4 \pm 0.2\%$, respectively.

To test the reliability of the sequential depletion procedure, we first performed reproducibility tests by analyzing five technical replicates (obtained from five independent depletions of the same serum), and also three biological replicates (obtained from three independent sera). For these analyses, proteins in the depleted samples were fractionated using C18 NuTips

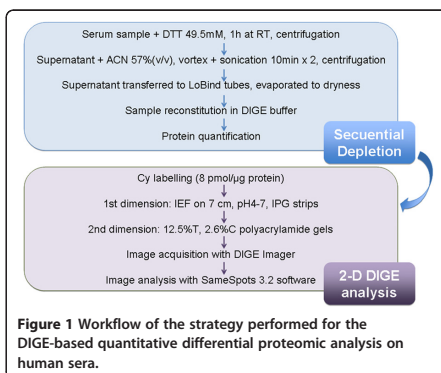


Figure 1 Workflow of the strategy performed for the DIGE-based quantitative differential proteomic analysis on human sera.

(Glygen, Columbia, MD, USA), by eluting with different concentrations of ACN (from 4% to 60%). By these means, ten fractions were obtained from each sample, which were then spotted onto a MALDI plate and analyzed by MS. Figure 2 demonstrates the high technical (A) and biological (B) reproducibility that we found on the sequentially depleted samples, which are also illustrated in the gel shown on Figure 2C.

Proteins in the crude (ND) and depleted (D) samples were separated by one- and two-dimensional gel electrophoresis, in order to evaluate the protein profiles of these two kinds of samples. The different profiles that were obtained after the sequential depletion are illustrated in Figure 3, showing the decrease of high molecular weight proteins that is characteristic of ACN-depleted sera, and also the enrichment of some proteins that were masked by albumin in the crude samples. As shown in the Figure, with this method we were able to deplete in an efficient way serum albumin (ALBU) and other high abundant proteins such as alpha-2-macroglobulin (A2MG), transferrin (TRFE), Complement 3α (C3α), or the heavy chain of IgG (IGGγ).

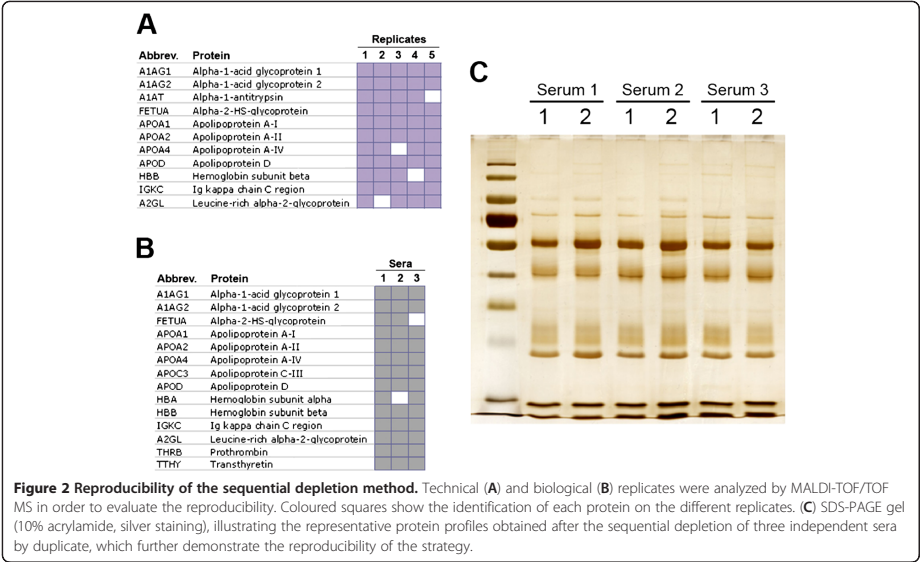
We randomly picked 54 spots from these 2D gels and identified the protein forms present in each of them, which correspond to 16 different proteins (Additional file 1: Figure S1 and Additional file 2: Table S1). Although many of them belong to the most abundant

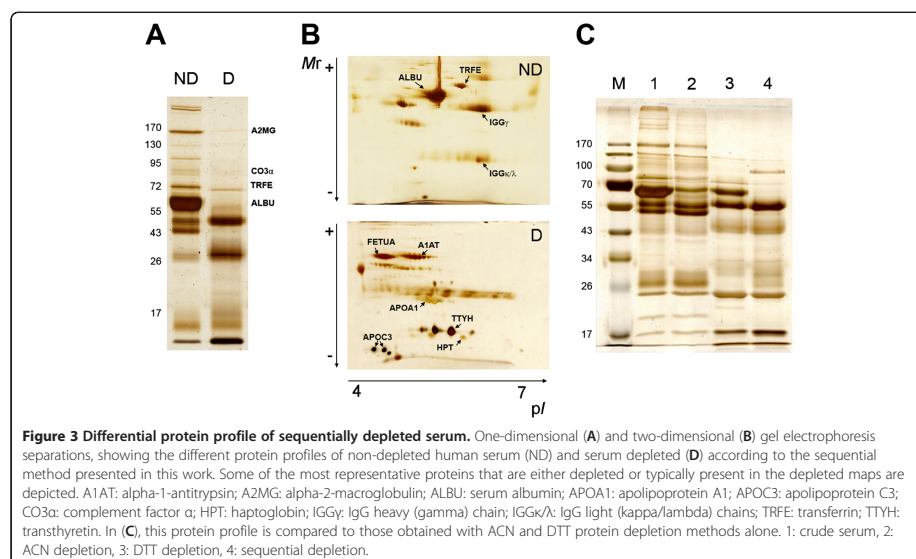
plasma proteins, as expected for a gel-based technique (which has an important bias towards abundant proteins), we were also able to identify two proteins that are reported to be present at less than 1 μmol/L in plasma, such as serum amyloid A (SAA1) or angiotensinogen (ANGT) [15]. These results support the usefulness of the strategy for the study of medium-abundant serum proteins.

2D-DIGE quantitative proteomics analysis on the sequentially depleted serum samples

The sequential depletion method was then applied to study the osteoarthritis-dependent modulation of proteins in serum. Those proteins that are altered in samples from diseased patients might be useful as OA biomarkers for diagnostic or prognostic purposes, and also for the evaluation of alternative therapies.

With this aim, serum samples from 40 patients (20 suffering severe OA and 20 healthy controls) were grouped into four pools of 10 samples each, to reduce the individual and biological variability that has been reported in plasma proteins [16]. Although this involves the loss of each sample individual data, this pooling strategy has demonstrated its usefulness for biomarker discovery in shotgun proteomics approaches [17]. Table 1 summarizes the characteristics of each pool. To perform the differential quantitative analysis,





we followed the DIGE strategy [18], which is the best quantitative gel-based approach up to date. It permits direct quantitative evaluation of changes, reduces inter-gel variation and false positives [9,19] and therefore results in highly confident data with biological significance. Moreover, one of the advantages of 2D-DIGE versus non-gel based quantitative proteomics techniques is that it can directly detect not only changes of protein quantity, but also posttranslational modifications of the

proteins [20], which are known to be remarkably abundant in serum samples.

The four serum pools were labelled with Cy dyes as described in Table 2. A technical replicate was performed with each sample by swapping the dye (that is, the same sample was labelled once with Cy3 and also with Cy5), in order to avoid labelling artefacts. Sample pairs were combined with the Cy2-labeled internal standard, and proteins were resolved by four-plex 2D gel electrophoresis. The 12 individual images obtained in this experiment (three from each gel, corresponding to Cy2-, Cy3- and Cy5-labeled samples) were analyzed with the SameSpots software, which allowed the normalization of spot intensities and the calculation of abundance ratios and statistical power of the results. All analyzed images are illustrated in Figure 4. An average of 210 protein spots were detected and matched on the gel images. We considered changes within 95% confidence interval ($p < 0.05$) and standardized average spot volume ratios exceeding 1.3.

Table 1 Characteristics of the osteoarthritis (OA) patients and controls included in this study

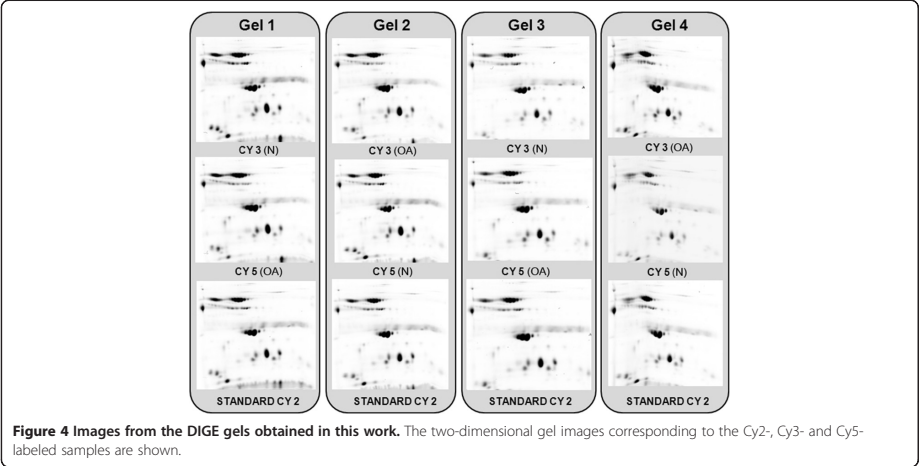
	Control	OA
Number of patients	20	20
Age ^{a)}	67.75 ± 9.08	75.85 ± 5.00
Protein ^{b)}	57.68 ± 5.06	56.38 ± 4.92
	Pool 1	Pool 2
Number of patients	10	10
Age	66.20 ± 8.83	76.10 ± 4.79
Protein	57.69 ± 5.15	54.69 ± 4.50
	Pool 3	Pool 4
Number of patients	10	10
Age	69.30 ± 9.52	75.60 ± 5.44
Protein	57.66 ± 5.25	58.07 ± 3.83

^{a)} Age shown as years with mean ± SD.

^{b)} Protein shown in µg/µl with mean ± SD.

Table 2 DIGE experimental design

Gel 1	Pool 1/Cy3	Pool 2/Cy5
Gel 2	Pool 1/Cy5	Pool 2/Cy3
Gel 3	Pool 3/Cy3	Pool 4/Cy5
Gel 4	Pool 3/Cy5	Pool 4/Cy3



Novel putative serum osteoarthritis biomarkers identified in this work

The analysis resulted in 46 spots significantly and reproducibly altered between OA and control samples (29 increased and 17 decreased), which are depicted on the 2D-DIGE gel image shown in Figure 5. For the identification of the altered proteins, spots were excised from the gels, digested in-gel and analysed by MALDI-MS/MS. Mascot database search using the MS/MS spectra

allowed the identification of the proteins present in all the spots. Altogether, these 46 spots correspond to 15 different protein forms that were identified as altered in OA. These proteins are listed in Table 3 with complete information comprising experimental and theoretical MW and pI values, accession number and identification parameters. A database search was carried out to assign them into different functional groups, which are also detailed in Table 3.

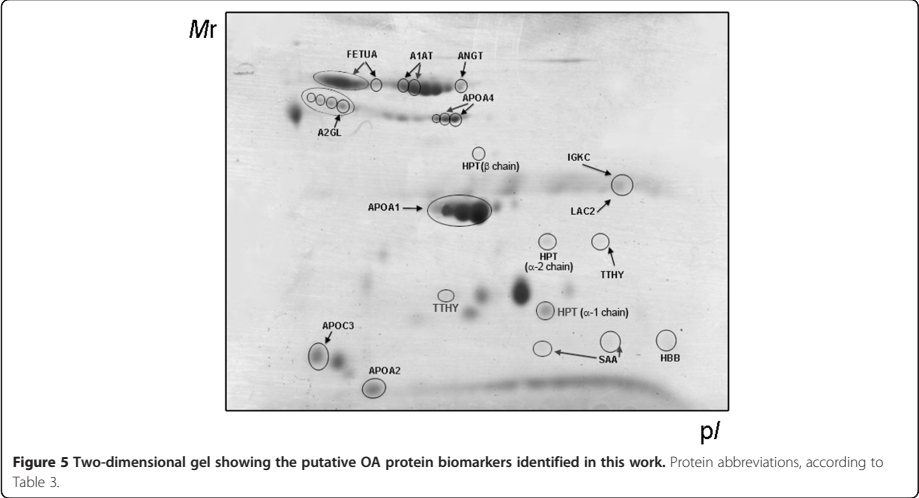


Table 3 Proteins identified in this work as altered in Osteoarthritis vs Control sera

Protein ID ^{a)}	Acc.No. ^{a)}	Name	Function	Ratio OA:N ^{b)}	Anova (p)	Predicted ^{c)}		Experimental ^{d)}		Score ^{e)}	Sequence tags ^{f)}
						MW (Da)	pI	MW (Da)	pI		
A1AT	P01009	Alpha-1-antitrypsin	Inhibitor of serine proteases	1.4	0.008	46878	5.37	40047	5.33	151	K. TDTSHHDDQDHPTFNKI K. ITPNLAEFAFSLYR.Q K. QINDYVEKGTQGKI K. GKWERPFVK.D K. DTEEDFHVDQVTVKV R. LGMFNIQHCKK K. LQHLENELTHDIITKF K. FLENEDRRS K. VFSNGADLSGVTEEAPLKL K. FNKPFVFLMIEQNTKS
FETUA	P02765	Alpha-2-HS-glycoprotein	Negative regulation of bone mineralization	2	0.004	40098	5.43	40098	5.43	102	K. CNLLAEKQYGCKA R. HTFMGVSLGSPSGEVSHPRK
ANGT	P01019	Essential component of the renin-angiotensin system (RAS)		1.3	0.001	53406	5.87	40056	5.48	28	Angiotensinogen K. ANAGKPKDPTFIPAPIQAKT
APOA1	P02647	Apolipoprotein A-I	Cholesterol metabolism	2.2	0.013	30759	5.56	31221	5.43	271	K. DSGRDYVSQFEGSALGK.Q K. LLDNWDVSTSTFSKL K. LLDNWDVSTSTFSKLRE K. VQPYLDDFQKK.W K. KWQEEMELYR.Q K. WQEEMELYR.Q K. VEPLRAELQEGAR.Q K. LSPLGEEMRDRA R. THLAPYSDLR.Q R. THLAPYSDLRQRL K. ATEHLSTLSEKAKPAEDLR.Q K. AKPAEDLR.Q
APOA2	P02652	Apolipoprotein A-II	Lipid transport	0.7	0.003	11282	6.26	11064	5.23	82	K. VKSPELQAEAKS K. SPELQAEAKS K. SKEQLTPLIKK
APOA4	P06727	Apolipoprotein A-IV	Lipid transport	0.6	0.005	45371	5.28	37738	5.47	37	K. LVPFATELHERL R. LEPYADQLRT R. TQVNTQAEQLRR.Q R. QLTPYAQR.M K. IDQNVEELKGR.L R. ISASAEELRQRL R. LAPLAEDVR.G R. LAPLAEDVRGNLR.G

Table 3 Proteins identified in this work as altered in Osteoarthritis vs Control sera (Continued)

											K. SLAELGGHLDQQVEEFRR
											K. SLAELGGHLDQQVEEFRR
											R. RVEPYGENFNKA
											K. ALVQQMEQLRQ
APOC3	P02656	Apolipoprotein C-III	Lipid degradation and transport	0.8	0.014	10846	5.23	13922	5.16	149	K. TAKDALSSVQESQVAQQAR.G
											K. DALSSVQESQVAQQAR.G
HPT	P00738	Haptoglobin (β chain)	Cellular iron ion homeostasis	2	0.007	45861	6.13	35377	5.74	90	K. GSFPWQAKM
											R. NANFKFTDHLKY
											K. YVMLPVADQDQCIRLH
											K. SPVGVQPILNEHTFCAGMSKY
HPT	P00738	Haptoglobin (α 1 chain)	Cellular iron ion homeostasis	1.4	0.006	45861	6.13	19281	5.98	68	R. YQCKNYYKL
											K. LRTEGDGWTLNNEKQ
											K. AVGDKLPECEAVCGKPKN
HPT	P00738	Haptoglobin (α 2 chain)	Cellular iron ion homeostasis	0.5	0.0002	45861	6.13	28110	5.97	135	R. TEGDGWYTLNDKKQ
											K. LRTEGDGWTLNNEKQ
											K. AVGDKLPECEAVCGKPKN
											K. AVGDKLPECEAVCGKPKN
HBB	P68871	Hemoglobin subunit beta	Involved in oxygen transport from the lung to peripheral tissues	1.6	0.014	16102	6.75	16102	6.75	43	R. LLVYPWTQRF
IGKC	P01834	Ig kappa chain C region	Immune response	0.6	0.003	11773	5.58	33273	6.42	107	- TVAAPSVFIFPPSDEQLKS
											K. SGTASWCLLNIFYPRE
LAC2	P0CG05	Ig lambda-2 chain C regions	Complement activation, classical pathway	0.6	0.003	11458	6.92	33273	6.42	72	K. AAPSVTLFPPSSEELQANKA
											K. QSNKNKYAASSYLSLTPEQWKS
											R. SYSCQVTHEGSTVEKT
											R. SYSCQVTHEGSTVEKTVAPTECS.-
A2GL	P02750	Leucine-rich alpha-2-glycoprotein	Protein binding (secreted)	1.5	0.005	38382	6.45	38507	5.47	88	R. YLFLNGNKLARV
											R. VAAGAFQGLRQ
											R. CAGPEAVKGQTLTLLAVAKS
SAA	P02735	Serum amyloid A protein	Acute-phase response	0.8	0.014	13581	6.28	15299	6.53	52	REANYIGSDKYFHAR.G
											R.FFGHGAEDSLADQAAANEWGR.S
TTHY	P02766	Transthyretin	Thyroid hormone-binding protein	2.2	0.022	15991	5.52	28287	6.3	136	R.KAADDTWEPFASGKT
											KAADDTWEPFASGKT

^{a)} Protein ID and accession number according to SwissProt database.

^{b)} Average volume ratio OAN, quantified by SameSpots software.

^{c)} Predicted molecular weight (MW) and Isoelectric point (pI) according to protein sequence and SwissProt database.

^{d)} Experimental molecular weight (MW) and Isoelectric point (pI) calculated by analysis of the gel images with SameSpots software.

^{e)} MS protein score obtained from MALDI-TOF/TOF spectra using Mascot search engine.

^{f)} Sequence tags identified by MALDI-TOF/TOF MS/MS.

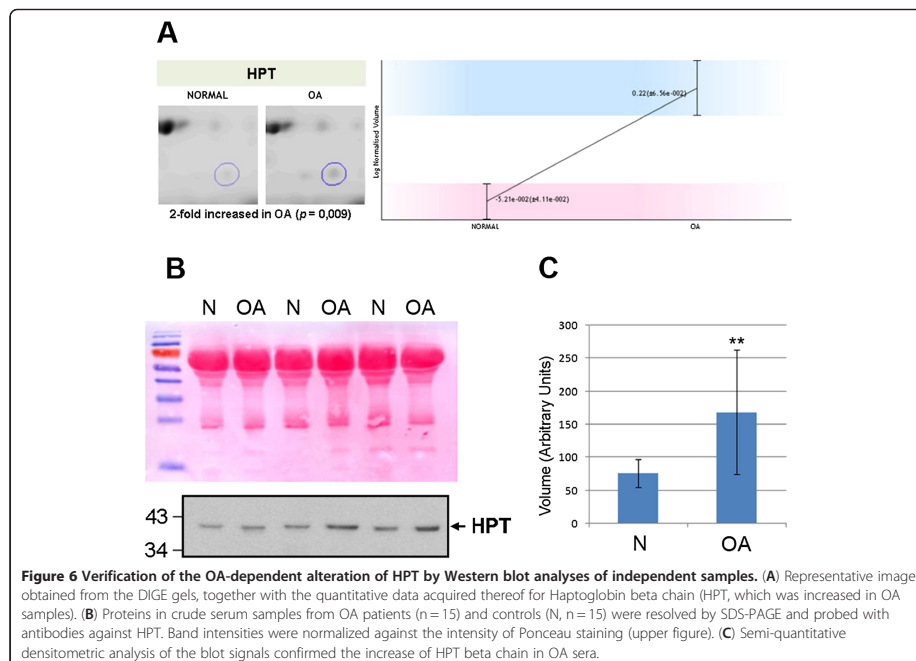
Interestingly, this altered panel differs substantially from that obtained in a previous work from our group based on serum immunodepletion and subsequent LC-MS/MS analysis [21]. In that work, several relatively high abundant proteins (such as complement components) were found increased in the sera from OA patients. On the contrary, most complement proteins are removed by the different depletion approach in the present study, while several apolipoproteins are identified. This demonstrates the complementarity of different depletion steps. Furthermore, although the gel-based strategy performed in this study did not allowed the identification of less abundant proteins, it demonstrates its usefulness of for characterizing disease-specific patterns of protein modifications, which are impossible to obtain from peptide/MS-based strategies. This includes not only the identification of a specific isoform of HPT beta chain as increased in OA, but also different forms of serum amyloid A, transthyretin or some apolipoproteins.

Nevertheless, both proteomic approaches led to the identification of several ubiquitous proteins related with lipid transport, immune response or protein binding. Although the molecules that had been employed typically as OA biomarkers are proteins directly or indirectly

involved in cartilage degradation, or proteins synthesized in an attempt at cartilage repair (including different type II collagen fragments or other cartilage extracellular matrix components) [22], increasing evidences suggest that the most promising strategy in OA would be the combination of different biomarkers into panels. These might include proteins highly specific to the pathology, but also ever-present proteins that can be altered during the progression of several unspecific processes such as inflammation, immune response, cellular death/proliferation, etc. The possible biomarker value of these proteins is exemplified by a recent work unravelling the key role of complement dysregulation in OA pathogenesis [23]. Furthermore, the identification in the present study of several proteins related with the lipid metabolism might be relevant for the definition of a new phenotype termed 'metabolic OA', which is recently acquiring more attention in the research community [24].

Osteoarthritis patients display an altered haptoglobin protein profile in serum

Interestingly, we detected three spots corresponding to the same protein accession number that exhibited different alterations in OA sera when compared to control



(two increased and the other decreased, with 1.5 to 2-fold variations). They were all identified as haptoglobin precursor (HPT, P00738) in the Mascot search, and the spot that was detected as increased in OA with a MW of around 35 kDa is illustrated in Figure 6A. HPT precursor is an N-linked glycoprotein [25] known to be cleaved into two chains (alpha and beta, see Additional file 3: Figure S2), which combine with free plasma hemoglobin and contribute to the maintenance of cellular iron homeostasis. Previous proteomic studies on synovial fluid, also performed using 2D-gels, reported remarkable interindividual differences in the haptoglobin patterns of the studied patients [26,27]. In the present study, however, we did not analyze individual samples, but pools made from 10 patients each. This strategy has proved its usefulness in our previous iTRAQ-based approach [21], and in the present work it has allowed the detection of changes in HPT chains with a high statistic significance. Nevertheless, we pursued also to verify the alteration of HPT protein pattern in OA sera by an orthogonal strategy, performing Western blot analyses on 30 new individual samples from OA and control donors ($n = 15$ from each condition). As shown in Figure 6B, a remarkable increase of the HPT beta chain (detected with a molecular weight around 35–40 kDa) can be detected in OA sera. We relatively quantified this increase as 2.22-fold by densitometric analysis of the blots, $p = 0.017$ (Figure 6C). Although the relationship of HPT polymorphisms with diseases such as cancer or rheumatoid arthritis has been largely studied due to the antioxidant [28], anti-inflammatory [29] and immunomodulatory properties [30] of this protein, this is the first time an alteration of HPT protein pattern is reported in OA.

Conclusions

In summary, we show in this work the application of a novel sequential depletion procedure for the search of osteoarthritis biomarkers in serum. This strategy has been coupled to a DIGE-based quantitative proteomic analysis, in order to find protein isoforms or fragments that are specific of the disease. By these means, we were able to identify 16 protein forms altered in the disease. Among these, we verified for the first time the OA-dependent increase of an haptoglobin chain. These data demonstrate the usefulness of the approach for protein biomarker discovery, and provide a list of potential OA biomarker candidates that might be subject of further validation studies.

Methods

Serum samples

The sera used for this study were obtained from OA patients and controls with no history of joint disease, which were characterized radiographically. Inclusion

criteria are fully described in a previous publication from our group [31]. The patients meeting these criteria were diagnosed with OA according to the American College of Rheumatology (ACR) criteria, and knee and hip radiographs from the participants were classified from grade I to grade IV according to the Kellgren and Lawrence (K/L) scoring system [32]. The patients were of both genders and ages ranged from 58 to 90 years. A population of 20 samples from K/L grade IV and 20 non-symptomatic controls was utilized for this study. Prior to proteomic analysis, the serum samples were grouped into four pools of 10 samples each to reduce individual and biological variability. Table 1 summarizes the characteristics of each pool.

Sequential depletion of high abundant proteins in serum

The sera were subjected to a sequential depletion protocol involving two precipitation steps (Figure 1): protein depletion was first performed with DTT according to the protocol described by Warder et al. [7] with minor modifications, and then further depleted with ACN according to Kay et al. [6]. With this aim, 2.2 μ l of 500 mM DTT were added to 20 μ l of pooled serum and incubated for 1 h at room temperature until a viscous precipitate persisted. This precipitate was pelleted by centrifugation 2×20 min at $14000 \times g$. Subsequently, the supernatant was transferred to a clean LoBind tube and diluted with water prior to the addition of 57% ACN [8]. Samples were briefly vortexed and sonicated twice for 10 min. Then, ACN-precipitated proteins were pelleted during 10 min at $14000 \times g$. Finally, the supernatant was transferred to clean LoBind tubes and evaporated to dryness in a Savant SPD121P SpeedVac (Thermo, Waltham, USA) for further analysis.

Protein quantification

Precipitated proteins were solubilized in an isoelectric focusing-compatible lysis buffer containing 8 M urea, 2 M thiourea, 30 mM Tris and 4% CHAPS. The proteins present in the pools were quantified ($n = 5$) using the Bradford method (modified with HCl) [33]. A BSA standard curve (0 to 4 mg/ml) and the samples were analyzed in triplicate by reading at 570 nm in a microplate reader (Multiskan[®] Plus, Labsystems, Quesada, Argentina).

Protein labeling and two-dimensional electrophoresis

The proteomics comparison between osteoarthritic and control sera was performed across four DIGE gels using 20 μ g of total protein fraction per CyDye[™] gel and two biological replicates for each condition. Proteins in each sample were fluorescently tagged with a set of matched fluorescent dyes according to the manufacturer's protocol for minimal labeling. To ensure that there were no

dye-specific labeling artifacts, Cy3 and Cy5 labels were swapped between two technical replicates of the same sample, whereas the pooled standard sample was labeled with Cy2. The standard pool was prepared by pooling 20 µg of proteins from each sample prior to labeling. In every case, 160 pmol of dye was used for 20 µg of proteins (8 pmol/µg). Labeling was performed for 30 min on ice in darkness, and the reaction was quenched with 1 µl of 10 mM L-lysine for 10 min under the same conditions.

The four pairs of Cy3- and Cy5-labeled samples (each containing 20 µg of protein) were combined and mixed with a 20-µg aliquot of the Cy2-labeled standard pool. The mixtures, containing 60 µg of protein, were diluted to 125 µl with rehydration buffer (7 M urea, 2 M thio-urea, 4% CHAPS, 0.002% bromophenol blue, 2% ampholytes (pH 3–11), and 97 mM Destreak reagent). Samples were loaded onto IPG strips (7 cm, pH 4–7 non-linear) by passive overnight rehydration. Isoelectric focusing was carried out on a IPGphor II IEF system (GE Healthcare) using the following conditions: 333 Vhr at 500 V, 40 min at 1000 V and finally 9166 Vhr at 5000 V. Prior to the second dimension run, the strips were equilibrated first for 15 min in equilibration buffer (100 mM Tris-HCl (pH 8.0), 6 M urea, 30% glycerol, and 2% SDS) with 2% DTT and then for another 15 min in the same buffer supplemented with 2.5% iodoacetamide and 0.002% bromophenol blue. The equilibrated strips were transferred onto 12.5% homogenous polyacrylamide gels (2.6% C) casted in low fluorescence glass plates. Electrophoresis was run at 150 V at 20°C.

Image acquisition and DIGE data analysis

The differentially labeled co-resolved proteins within each gel were imaged at 100 dots/inch resolution using a DIGE Imager scanner (GE Healthcare). Cy2-, Cy3-, and Cy5-labeled images of each gel were acquired at excitation/emission values of 488/520, 523/580, and 633/670 nm, respectively. Gels were scanned directly between the glass plates, and the 16-bit image file format images were exported for data analysis. After imaging for CyDyes, the gels were removed from the plates and stained with colloidal Coomassie following standard procedures.

Semi-automated image analysis was performed with Progenesis SameSpots V3.2 software (Nonlinear Dynamics). Multiplexed analysis was selected for DIGE experiments and a representative gel image was chosen as reference. Spots were detected and their normalized volumes were ranked on the basis of ANOVA p-values and fold changes.

Ultrasonic in-gel enzymatic digestion

The gel spots of interest were manually excised and transferred to microcentrifuge tubes. Samples selected

for analysis were subjected to ultrasonic in-gel enzymatic digestion, according to the ultrafast proteolytic digestion protocol previously described [34,35]. Briefly, protein bands were washed with water and acetonitrile, reduced with DTT and alkylated with IAA in an ultrasonic bath (Sonorex RK 31 H, Bandelin, Berlin, Germany) operating at 35 kHz (100% amplitude). Then, the gel was washed again, dried and trypsin (375 ng in 25 µl) was added to the dried protein bands. The in-gel protein digestion was performed in a sonicator model SONOPULS HD 2200 (Bandelin, Berlin, Germany) operating at 50% amplitude for 2 min. Finally, supernatants were collected and further peptide extraction was performed on the remaining gel piece with acetonitrile 50% V/V, trifluoroacetic acid 0.1%V/V during 2 min at 50% amplitude in the sonicator. Supernatants were pooled, evaporated to dryness and reconstituted in 5 µl of formic acid 0.3% V/V.

Mass spectrometry (MS) analysis

The samples were analyzed using the Matrix-assisted laser desorption/ionization (MALDI)-Time of Flight (TOF)/TOF mass spectrometer 4800 Proteomics Analyzer (ABSCIEX, MA, USA) and 4000 Series Explorer™ Software (ABSCIEX). Data Explorer version 4.2 (ABSCIEX) was used for spectra analyses and generating peak-picking lists. All mass spectra were internally calibrated using autoprolytic trypsin fragments and externally calibrated using a standard peptide mixture (Sigma-Aldrich). TOF/TOF fragmentation spectra were acquired by selecting the 20 most abundant ions of each MALDI-TOF peptide mass map (excluding trypsin autolytic peptides and other known background ions). An average of 2000 laser shots were employed per fragmentation spectrum.

Database search

The amino acid sequence tags obtained from each peptide fragmentation in MS/MS analyses were used to search for protein candidates using Mascot version 2.2 from Matrix Science (<http://www.matrixscience.com>). Peak intensity was used to select up to 50 peaks per precursor for MS/MS identification. Tryptic autolytic fragments, keratin and matrix-derived peaks were removed from the dataset used for the database search. The searches for peptide mass fingerprints and tandem MS spectra were performed in the SwissProt knowledgebase (2011_05 release version, June 2011), by searching in the UniProtKB/Swiss-Prot (<http://www.expasy.ch/sprot>) database, containing 528048 sequences and 186939477 residues, with taxonomy restriction (*Homo sapiens*, 20239 sequences). Fixed and variable modifications were considered (Cys as S-carbamidomethyl derivative and Met as oxidized methionine, respectively), allowing one trypsin missed cleavage site and a mass tolerance of 50 ppm. MS

identifications were accepted as positive when at least five peptides matched and at least 20% of the peptide coverage of the theoretical sequences matched within a mass accuracy of 50 or 25 ppm with internal calibration. For MS/MS identifications, a precursor tolerance of 50 ppm and MS/MS fragment tolerance of 0.3 Da were used. Probability scores were significant at $p < 0.01$ for all matches.

Western blot analysis

One-dimensional Western blot tests were performed according to standard procedures. Briefly, 40 µg of serum proteins were loaded and resolved on standard 15% polyacrylamide SDS-PAGE gels. Separated proteins were then electroblotted onto PVDF membranes (Immobilon P, Millipore, Bedford, MA). Equivalent loadings were verified by Ponceau Red staining after transference. Membranes were blocked in Tris-buffered saline (pH 7.4) containing 0.1% Tween 20 (TBST) and 5% non-fat dried milk for 60 min at room temperature. The blots were then hybridized overnight at 4°C with monoclonal antibodies against the beta chain of HPT (ab13429, 1:500 dilution, Abcam, Cambridge, UK), diluted in TBST with 2% nonfat milk. After thorough washing with TBST, immunoreactive bands were detected by chemiluminescence using the corresponding horseradish peroxidase-conjugated secondary antibodies and ECL detection reagents (GE Healthcare), and then digitized using an LAS 3000 image analyzer. Quantitative changes in band intensities were evaluated with ImageQuant 5.2 software (GE Healthcare). The densitometry values of the Western blot bands containing the HPT were normalized against those of Ponceau staining obtained from the same membranes. Then, the relative abundance of HPT was calculated by obtaining the ratio of the normalized densitometric values between normal and OA samples. Statistical p values of the densitometry data were obtained by application of Mann–Whitney U test using SPSS version 15.0 program.

Additional files

Additional file 1: Figure S1. Representative 2D map of the sequentially depleted serum, with the protein identifications that were carried out. Identifications are listed in Additional file 2: Table S1.

Additional file 2: Table S1. Mass spectrometry data of the protein identifications performed in the present study. Spot numbers, according to Additional file 1: Figure S1.

Additional file 3: Figure S2. Amino acid sequence of Haptoglobin, with the MS data corresponding to the identification of alpha and beta chains.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CF carried out the main experimental work, and prepared the figures and tables. VC participated in the DIGE experiments and gel bioinformatic analyses. PFP collected and classified the serum samples. JLCM participated in study design and helped with the data analysis. CRR participated in the study design, the interpretation of the data and drafted the manuscript. FJB conceived and coordinated the project and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We are grateful to Natividad Oreiro and Carlos Fernández, from the Rheumatology Division of Hospital Universitario A Coruña, for providing serum samples. This study was supported by grants from Fondo Investigación Sanitaria-Spain (CIBER-CB06/01/0040; PI08/2028; PI11/2397); and Secretaría I + D + I Xunta de Galicia (09CSA043383PR, 10CSA916058PR). C. F. and V. C. are supported by Foundation of the Hospital Universitario A Coruña. J.-L. C.-M. is grateful to the Xunta de Galicia (Spain) for the Isidro Parga Pondal program. C.R.-R. is supported by Fondo Investigación Sanitaria-Spain (CP09/00114).

Author details

¹Rheumatology Division, ProteoRed/ISCIII Proteomics Group, INIBIC – Hospital Universitario de A Coruña, 15006 A Coruña, Spain. ²REQUIMITE, Departamento de Química, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, Caparica, Portugal. ³BIOSCOPE Group, Physical Chemistry Department, Science Faculty, University of Vigo at Ourense Campus, Ourense, Spain. ⁴CIBER-BBN Instituto de Salud Carlos III, A Coruña, Spain.

Received: 13 February 2012 Accepted: 6 September 2012

Published: 12 September 2012

References

- Anderson NL, Anderson NG: The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 2002, **1**(11):845–867.
- Roche S, Tiers L, Provansal M, Piva MT, Lehmann S: Interest of major serum protein removal for Surface-Enhanced Laser Desorption/Ionization - Time Of Flight (SELDI-TOF) proteomic blood profiling. *In: Proteome Sci* 2006, **4**:20.
- Echan LA, Tang HY, Ali-Khan N, Lee K, Speicher DW: Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* 2005, **5**(13):3292–3303.
- Righetti PG, Castagna A, Antoniolli P, Boschetti E: Prefractionation techniques in proteome analysis: the mining tools of the third millennium. *Electrophoresis* 2005, **26**(2):297–319.
- Roche S, Tiers L, Provansal M, Severino M, Piva MT, Jouin P, Lehmann S: Depletion of one, six, twelve or twenty major blood proteins before proteomic analysis: the more the better? *In: J Proteomics*. vol. 72. Netherlands 2009, **72**:945–951.
- Kay R, Barton C, Ratcliffe L, Matharoo-Ball B, Brown P, Roberts J, Teale P, Creaser C: Enrichment of low molecular weight serum proteins using acetonitrile precipitation for mass spectrometry based proteomic analysis. *Rapid Commun Mass Spectrom* 2008, **22**(20):3255–3260.
- Wardner SE, Tucker LA, Streltzer TJ, McKeegan EM, Meuth JL, Jung PM, Saraf A, Singh B, Lai-Zhang J, Gagne G, et al: Reducing agent-mediated precipitation of high-abundance plasma proteins. *In: Anal Biochem*. vol. 387. United States 2009, **387**:184–193.
- Fernández C, Santos HM, Ruiz-Romero C, Blanco FJ, Capelo JL: A comparison of depletion vs. equalization for reducing high abundance proteins in human serum. *Electrophoresis*; 2011.
- Marouga R, David S, Hawkins E: The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal Bioanal Chem* 2005, **382**(3):669–678.
- Pritzker K, In: Osteoarthritis: Pathology of osteoarthritis. New York: Oxford University Press; Edited by Brandt K, Doherty M, Lohmander L; 1998:50–61.
- Woolf A, Pfleger B: Burden of major musculoskeletal conditions. *Bulletin of the World Health Organization* 2003, **81**:10.
- Peat G, Thomas E, Duncan R, Wood L, Wilkie R, Hill J, Hay EM, Croft P: Estimating the probability of radiographic osteoarthritis in the older patient with knee pain. *Arthritis Rheum* 2007, **57**(5):794–802.

13. Kraus VB, Burnett B, Coindreau J, Cottrell S, Eyre D, Gendreau M, Gardiner J, Garner P, Hardin J, Henrotin Y, et al: **Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis.** *Osteoarthritis Cartilage* 2011, **19**(5):515–542.
14. Ruiz-Romero C, Blanco FJ, In: *Osteoarthritis Cartilage: Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis*. vol. 18. England: 2010 Osteoarthritis Research Society International: Published by Elsevier Ltd; 2010:500–509.
15. Hortin GL, Sviridov D, Anderson NL: **High-abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance.** *Clin Chem* 2008, **54**(10):1608–1616.
16. Nedelkov D, Kiernan UA, Niederkofer EE, Tubbs KA, Nelson RW: **Investigating diversity in human plasma proteins.** *Proc Natl Acad Sci U S A* 2005, **102**(31):10852–10857.
17. Alsaif M, Guest PC, Schwarz E, Reif A, Kittel-Schneider S, Spain M, Rahmouni H, Bahn S: **Analysis of serum and plasma identifies differences in molecular coverage, measurement variability, and candidate biomarker selection.** *Proteomics Clin Appl* 2012, **6**(5–6):297–303.
18. Unlu M, Morgan ME, Minden JS: **Difference gel electrophoresis: a single gel method for detecting changes in protein extracts.** *Electrophoresis* 1997, **18**(11):2071–2077.
19. Lilley KS, Friedman DB: **All about DIGE: quantification technology for differential-display 2D-gel proteomics.** *Expert Rev Proteomics* 2004, **1**(4):401–409.
20. Minden J: **Comparative proteomics and difference gel electrophoresis.** *Biotechniques* 2007, **43**(6):739–741. 743 passim.
21. Fernandez-Puente P, Mateos J, Fernandez-Costa C, Oreiro N, Fernandez-Lopez C, Ruiz-Romero C, Blanco FJ: **Identification of a panel of novel serum osteoarthritis biomarkers.** *J Proteome Res* 2011, **10**(11):5095–5101.
22. Olsen AK, Sondergaard BC, Byrjalsen I, Tanko LB, Christiansen C, Muller A, Hein GE, Karsdal MA, Qvist P: **Anabolic and catabolic function of chondrocyte ex vivo is reflected by the metabolic processing of type II collagen.** *Osteoarthritis Cartilage* 2007, **15**(3):335–342.
23. Wang Q, Rozelle AL, Lepus CM, Scanzello CR, Song JJ, Larsen DM, Crish JF, Bebek G, Ritter SY, Lindstrom TM, et al: **Identification of a central role for complement in osteoarthritis.** *In: Nat Med.* vol. 17. United States 2011, **17**:1674–1679.
24. Blanco FJ, Ruiz-Romero C: **Osteoarthritis: Metabolomic characterization of metabolic phenotypes in OA.** *Nat Rev Rheumatol* 2012, **8**(3):130–132.
25. Zhang H, Li XJ, Martin DB, Aebersold R: **Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry.** *In: Nat Biotechnol.* vol. 21. United States 2003, **21**:660–666.
26. Sinz A, Bantscheff M, Mikkat S, Ringel B, Drynda S, Kekow J, Thiesen HJ, Glocker MO: **Mass spectrometric proteome analyses of synovial fluids and plasmas from patients suffering from rheumatoid arthritis and comparison to reactive arthritis or osteoarthritis.** *Electrophoresis* 2002, **23**(19):3445–3456.
27. Yamagiwa H, Sarkar G, Charlesworth MC, McCormick DJ, Bolander ME: **Two-dimensional gel electrophoresis of synovial fluid: method for detecting candidate protein markers for osteoarthritis.** *J Orthop Sci* 2003, **8**(4):482–490.
28. Gutteridge JM: **The antioxidant activity of haptoglobin towards haemoglobin-stimulated lipid peroxidation.** *Biochim Biophys Acta* 1987, **917**(2):219–223.
29. Jue DM, Shim BS, Kang YS: **Inhibition of prostaglandin synthase activity of sheep seminal vesicular gland by human serum haptoglobin.** *Mol Cell Biochem* 1983, **51**(2):141–147.
30. Arredouani M, Matthijs P, Van Hoeyveld E, Kasran A, Baumann H, Ceuppens JL, Stevens E: **Haptoglobin directly affects T cells and suppresses T helper cell type 2 cytokine release.** *Immunology* 2003, **108**(2):144–151.
31. Rego I, Fernandez-Moreno M, Fernandez-Lopez C, Gomez-Reino JJ, Gonzalez A, Arenas J, Blanco FJ: **Role of European mitochondrial DNA haplogroups in the prevalence of hip osteoarthritis in Galicia, Northern Spain.** *In: Ann Rheum Dis.* vol. 69. England 2010, **69**:210–213.
32. Kessler S, Guenther KP, Puhl W: **Scoring prevalence and severity in gonarthrosis: the suitability of the Kellgren & Lawrence scale.** *Clin Rheumatol* 1998, **17**(3):205–209.
33. Kruger NJ: **The Bradford method for protein quantitation.** *Methods Mol Biol* 1994, **32**:9–15.
34. Rial-Otero R, Carreira RJ, Cordeiro FM, Moro AJ, Fernandes L, Moura I, Capelo JL: **Sonoreactor-based technology for fast high-throughput proteolytic digestion of proteins.** *J Proteome Res* 2007, **6**(2):909–912.
35. Galesio M, Vieira DV, Rial-Otero R, Lodeiro C, Moura I, Capelo JL: **Influence of the protein staining in the fast ultrasonic sample treatment for protein identification through peptide mass fingerprint and matrix-assisted laser desorption ionization time of flight mass spectrometry.** *J Proteome Res* 2008, **7**(5):2097–2106.

doi:10.1186/1477-5956-10-55

Cite this article as: Fernández-Costa et al.: Sequential depletion of human serum for the search of osteoarthritis biomarkers. *Proteome Science* 2012 **10**:55.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



3.3. Artículo III: Evaluación de los ultrasonidos indirectos de alta intensidad en la digestión de proteomas complejos con tripsina inmovilizada

Recientemente se están empleando los ultrasonidos de alta intensidad para acelerar tratamientos de muestra. En el campo de la proteómica, se utilizan como una herramienta para llevar a cabo la digestión de proteomas complejos en menos de 5 min. Hasta el momento se ha descrito el uso de los ultrasonidos en la digestión de proteínas con tripsina en solución o tripsina inmovilizada en soportes sólidos. Las ventajas que presenta el empleo de la tripsina inmovilizada son, entre otras, la posibilidad de reutilizar el enzima y de emplear una mayor cantidad de enzima sin comprometer el análisis por espectrometría de masas, debido a la presencia de picos de tripsina interferentes. La aplicación de los ultrasonidos en la digestión con tripsina inmovilizada se llevó a cabo con una sonda de ultrasonidos directamente dentro de la muestra líquida. Investigaciones recientes en la aplicación de los ultrasonidos a la proteómica aconsejan emplear ultrasonidos de alta intensidad indirectos (IHI-UE). Los IHI-UE permiten tratar más muestras a la vez, evitan la contaminación de las muestras proveniente de la punta de la sonda, permiten la refrigeración de las muestras y el tratamiento sellado para muestras peligrosas.

El objetivo de este trabajo fue estudiar la eficacia de la digestión con tripsina inmovilizada acelerada por ultrasonidos indirectos en muestras de suero humano. Para ello, las muestras de suero se digirieron con tripsina inmovilizada en bolitas magnéticas o con tripsina en disolución. Ambas reacciones se aceleraron con ultrasonidos tanto directos como indirectos. La eficacia de los dos métodos se comparó mediante marcaje isotópico con ^{18}O .

La eficacia de la tripsina inmovilizada se estudió mediante un experimento en el que tres sueros distintos se deplecionaron con el protocolo de depleción secuencial desarrollado por nuestro grupo, posteriormente se redujeron y alquilaron las proteínas. Cada muestra se dividió en dos partes iguales, una fracción se digirió con tripsina inmovilizada mientras que la otra fracción se digirió con tripsina en disolución, ambas bajo el efecto de los ultrasonidos indirectos. Las muestras digeridas se limpiaron y a continuación se realizó un marcaje con ^{18}O de la muestra digerida con tripsina en disolución, para posteriormente mezclar con la muestra sin

marcar y cuantificar en el mismo espectro de masas ambas muestras por cuadruplicado.

Se hizo un experimento para comparar los ultrasonidos directos e indirectos con la tripsina inmovilizada. Para ello, se deplecionaron cuatro sueros diferentes con la depleción secuencial por duplicado. Se redujeron y alquilaron las proteínas y cada muestra se dividió en dos partes iguales. Se digirieron con tripsina inmovilizada y una fracción se trató con ultrasonidos directos y la otra fracción con ultrasonidos indirectos. Las muestras digeridas se limpiaron y se marcó con ^{18}O la muestra tratada con ultrasonidos directos. Las muestras marcada y sin marcar se juntaron y se analizaron por triplicado por espectrometría de masas, obteniendo las cuantificaciones correspondientes.

Para la cuantificación se emplearon los péptidos comunes a la muestra marcada y sin marcar y aquellos péptidos con una eficacia superior al 90% de la doble incorporación del ^{18}O en el extremo carbono terminal (C-terminal).

En la comparación de la tripsina inmovilizada con la tripsina en disolución bajo el efecto de los ultrasonidos indirectos, el ratio se calculó dividiendo la intensidad de la muestra digerida con tripsina en solución entre la intensidad de la muestra digerida con tripsina inmovilizada. En la primera muestra encontramos 22 de 23 péptidos cuantificados, en la segunda muestra 11 de 16 péptidos y en la tercera 13 de 20 péptidos cuantificados, en los tres casos, con un ratio mayor a 1.2. También se encontró un considerable número de péptidos comunes a las dos condiciones que no se pudieron cuantificar en la mezcla por estar sólo en la condición marcada, esto fue debido a que la digestión en solución fue tan eficaz comparada con la tripsina inmovilizada que la intensidad de los péptidos no marcados no se detectaron.

En el estudio de la influencia del tipo de ultrasonidos aplicado con la tripsina inmovilizada en bolitas, se observó que con los ultrasonidos indirectos se formaron agregados de las bolitas impidiendo el reparto de la tripsina de forma homogénea por toda la muestra. No resultó así con los ultrasonidos directos, en donde se observó una mezcla homogénea. Estos efectos se muestran en dos vídeos del material suplementario del artículo. Los resultados del experimento realizado para

comparar estos dos tipos de sonicación mediante el marcaje con ^{18}O mostraron que en el 70% de los casos la digestión con ultrasonidos directos fue más eficiente, obteniendo ratios desde 1.2 hasta 4.4 veces mayor que con los ultrasonidos indirectos.



Contents lists available at SciVerse ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Short communication

An assessment of the indirect high intensity ultrasonic assisted cleavage of complex proteomes with immobilized trypsin.

C. Fernández-Costa^a, C. Ruiz-Romero^{a,b}, F.J. Blanco^{a,c}, H.M. Santos^d, J.L. Capelo^{d,*}^a Rheumatology Division, Proteomics Group-ProteoRed/ISCIII, INIBIC-CHUAC, A Coruña, Spain^b CIBER-BBN, INIBIC-CHUAC, A Coruña, Spain^c RIER-RED de Inflamación y Enfermedades Reumáticas, INIBIC-CHUAC, A Coruña, Spain^d REQUIMTE, Departamento de Química, BIOSCOPE Group, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, Caparica, Portugal

ARTICLE INFO

Article history:

Received 8 October 2012

Accepted 11 December 2012

Available online 22 December 2012

Keywords:

Protein cleavage

Biomarker discovery

Ultrasonic

Ultrasonid

Probe

Cup-horn

Immobilized trypsin

ABSTRACT

The indirect high intensity ultrasonic assisted cleavage of complex proteomes using immobilized trypsin has been assessed. It has been found that the formation of aggregates between the beds supporting the immobilized trypsin is promoted. This affects the efficiency of the digestion process, which can be 100 times lower than the digestion efficiency obtained with in-solution trypsin. Through the use of isotopic peptide labelling with ^{18}O , it has been quantified that the digestion efficiency over serum samples can be 1.2–100 times higher for the 70% of the peptides when indirect ultrasonication is replaced by direct ultrasonication. Therefore, direct high intensity ultrasonic assisted cleavage of proteins is proposed as an alternative to be combined with immobilized trypsin.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Quantification of proteins through the labelling of peptides with ^{18}O can be done following two different approaches [1–3]. In the first approach, the sample is digested in ^{18}O water. Therefore, the digestion of the protein and the labelling of the peptides take place at the same time. It has been demonstrated, however, that if the digestion of proteins is done at the same time that the peptide labelling, there is a lack in the double ^{18}O incorporation if this approach is used [4–7]. To solve this problem, some algorithms have been developed to correct this inefficient double labelling [8,9]. However, the simplest way to overcome this drawback is by doing the protein digestion and the protein labelling in two separated and consecutive steps. Thus, in the second approach, the sample is first digested in normal water, then dried and finally recomposed in ^{18}O water. This allows for a reproducible and efficient double incorporation of ^{18}O , leading to the accurate and precise quantification of proteins [3,10].

Recently, the use of ultrasonic energy, UE, in ^{18}O -based workflows for protein quantification [11,12], as a tool to speed the

sample treatment pipeline has been reported. Thus, when UE is used, digestion of complex proteomes, such as plasma or serum, can be done in less than 5 min [10,12]. However, the use of UE seems to speed only the mechanisms involved in the protein digestion, being negligible its influence in the efficiency or in the speed of the ^{18}O incorporation [6–12]. In other words, UE can only speed the protein digestion but not the ^{18}O labelling of peptides. Up to date, the use of UE in the digestion of proteins or complex proteomes has been described using trypsin immobilized in solid supports or trypsin in solution. The use of immobilized trypsin brings a number of benefits, such as (i) the possibility to reuse the enzyme, and (ii) the possibility to use higher amounts of enzyme without compromising the subsequent mass spectrometric analysis with the presence of interfering peptides [10,12]. Research dealing with UE and immobilized trypsin was developed using an ultrasonic probe to deliver the ultrasonic energy into the sample. Current research focused towards the application of ultrasonic energy in proteomics, has addressed the convenience to use indirect high intensity ultrasonic energy, IHI-IUE, through cup-horns instead of ultrasonic probes, as a better way to speed digestion of proteins [5]. IHI-UE allows to treat more samples at once, it avoids sample contamination and cross-contamination caused by the tip of the ultrasonic probe, it allows sample cooling and dangerous samples are treated sealed [10]. Recently, however, we have verified that IHI-UE is not an appropriate way to speed protein cleavage when immobilized trypsin is used. Unexpectedly, we have observed that the beds tend to form aggregates when they

*Corresponding author at: Departamento de Química-Física, Faculdade de Ciências, Campus de Ourense, Universidade de Vigo, 32004, Spain.
Tel.: +34 610 835 903.

E-mail addresses: jlcapelom@uvigo.es, jlcapelom@gmail.com, jlc@fct.unl.pt (J.L. Capelo).

URL: <http://www.bioscopegroup.org> (J.L. Capelo).

are submitted to IHI-UE, leading to a decrease in the digestion efficiency (see videos of supplementary material, VSM1 and VSM2).

Supplementary material related to this article can be found online at <http://dx.doi.org/10.1016/j.talanta.2012.12.022>.

In the present academic paper a complex proteome, human serum, was digested with trypsin immobilized in beds or with trypsin in solution. Both procedures were accelerated with UE provided by either a cup-horn or an ultrasonic probe. The efficiency of both methods of digestion was compared using ^{18}O isotopic labelling. The results clearly demonstrated that UE provided by a cup-horn in combination with immobilized trypsin has a worst performance than a cup-horn in combination with trypsin in solution.

2. Results and discussion

2.1. Materials and methods

2.1.1. Reagents

DL-dithiothreitol (DTT, $\geq 99\%$), iodoacetamide (IAA, $\geq 99\%$), trifluoroacetic acid (TFA, 99% FOR LC-MS), acetonitrile (ACN, LC-MS CHROMASOLV), water (LC-MS CHROMASOLV) and ammonium bicarbonate ($\geq 99\%$) were purchased from Sigma (Steinheim, Germany), and were used for protein precipitation, reduction and alkylation. Trypsin (sequencing grade) from Roche (Mannheim, Germany) and Mag-Trypsin (5% suspension) from Clontech (Mountain View, USA) were used for protein digestion. NuTips large 10–200 μL , C-18 for sample desalting were purchased from Glygen (Columbia, USA). Water- ^{18}O (97 atom % ^{18}O) from Sigma-Aldrich (Steinheim, Germany). α -Cyano-4-hydroxycinnamic acid puriss for MALDI-MS from Fluka (Steinheim, Germany) was used as MALDI matrix. Calibration 1, 4700 Proteomics Analyzer Calibration Mixture from Applied Biosystems was used as mass calibration standard for MALDI-TOF/TOF-MS measurements.

2.2. Sera samples

The human sera samples were obtained from anonymous donors in the Complejo Hospitalario Universitario de A Coruña, Spain. Osteoarthritis (OA) patients and controls were characterized radiographically. All patients signed the informed consent. The study was approved by the local ethics committee of Galicia (Spain).

Serum samples of OA and Normal patients were stored at -80°C until further processing.

2.3. Sequential protein depletion

Sequential protein depletion was done according to the protocol previously developed in our laboratory [13]. The sera were subjected to a sequential depletion protocol involving two precipitation steps. Protein depletion was first performed with DTT according to the protocol described by Warder et al. [14] and the second depletion was performed with ACN according to the protocol described by Kay et al. [15] with minor modifications. In brief, 2.2 μL of 500 mM DTT were added to 20 μL of serum and vortexed shortly. Then the sample was incubated at room temperature for 1 h until a viscous white precipitate was observed, then the sample was centrifuged at 14000g for 2×20 min. The supernatant was transferred into a clean LoBind tube and was further depleted with 57% (v/v) ACN, vortexed to mix and sonicated 10 min in an ultrasonic bath. Then the sample was vortexed again briefly and sonicated 10 min. The protein precipitate was then pelleted by centrifugation at 14000g for 10 min. The pellet was

discharged and the supernatant was transferred into a clean LoBind tube and evaporated to dryness in a vacuum concentrator centrifuge Savant SPD121P SpeedVac (Thermo, Waltham, USA) and stored at -20°C .

2.4. Serum digestion in homogeneous phase (trypsin in solution)

Ultrasonic in solution digestion was done according to the ultrafast proteolytic digestion protocol previously developed in our laboratory [4]. The pellet was re-suspended in 20 μL of ammonium bicarbonate (AmBi) 25 mM and then 10 μL of ACN were added. The sample was then vortexed and sonicated 1 min (50% amplitude) in the sonicator (SONOPULS HD 2200 with cup-horn BB6 accessory, Bandelin). Protein cysteine residues were reduced with 2 μL DTT 110 mM, and then the sample was vortexed for a short time and then sonicated again for 1 min (50% amplitude) in the sonicator. The resulting cysteines were then blocked with 2 μL IAA 600 mM and then the sample was vortexed for a short duration and then sonicated again for 1 min (50% amplitude) in the sonicator. Afterwards, 10 μL of DTT 110 mM were added for IAA inactivation. Then, the sample was diluted to a final volume of 200 μL with AmBi 12.5 mM. Afterwards, trypsin was added with ratio 1:20 (w/w) trypsin/protein, as recommended by the manufacturer. The digestion was performed in the sonicator operating at 50% amplitude for 2.5 min. Then, 2 μL of formic acid 50% (v/v) were added to stop the enzyme activity. Samples were acidified with 1 μL of pure TFA and desalting was then performed using NuTipTM pipette tips.

2.5. Serum digestion in heterogeneous phase (immobilized trypsin)

The pellet was resuspended in 20 μL of ammonium bicarbonate (AmBi) 25 mM and then 10 μL of ACN were added, the sample was vortexed and sonicated 1 min (50% ultrasonication amplitude) in the sonicator (SONOPULS HD 2200 with cup-horn BB6 accessory, from Bandelin, USA). Protein cysteine residues were reduced with the addition of 2 μL DTT 110 mM, vortexed and sonicated again 1 min (50% amplitude) in the sonicator. The resulting cysteines were then blocked with 2 μL IAA 600 mM and vortexed and sonicated again 1 min (50% amplitude) in the sonicator. 10 μL of DTT 110 mM were added for IAA inactivation. The sample was diluted to a final volume of 200 μL with AmBi 12.5 mM and then was evaporated to dryness in a vacuum concentrator centrifuge. Prior to use, 200 μL of Mag-trypsin (5% suspension) were washed with H_2O -LCMS three times. Then, the sample was reconstituted in this volume of clean immobilized trypsin and was digested in the sonicator with cup-horn at 50% of amplitude for 2.5 min. The supernatant was recovered employing a magnet for retaining the Mag-trypsin beds. The digested sample was acidified with 1 μL of pure TFA and was subsequently NuTipTM desalted. The cleaned sample was evaporated to dryness and label with ^{18}O as described in Section 2.7.

2.6. Digestion with immobilized trypsin assisted with ultrasonic probe

The depleted serum was reduced, alkylated and digested with the same protocol as described in Section 2.5, but the digestion was assisted with direct ultrasounds (Ultrasonic robe DRH-UP50H, Dr. Hielscher, 0.5 mm tip) at 50% of amplitude for 2.5 min. The supernatant was recovered employing a magnet for retaining the Mag-trypsin beds. The digested sample was acidified with 1 μL of pure TFA and was subsequently NuTipTM desalted. The cleaned sample was evaporated to dryness and label with ^{18}O as described in Section 2.7.

2.7. Trypsin-catalyzed ^{18}O labelling

The sample was re-composed in 10 μl of ammonium acetate 25 mM/ACN 10% v/v and was evaporated to dryness. Then, the sample digested with in-solution trypsin was reconstituted in 5 μl of H_2^{18}O with 5 μl of Mag-trypsin cleaned and dried. The sample digested with immobilized trypsin was reconstituted in 5 μl of H_2^{16}O in presence of Mag-trypsin. The samples were vortexed and centrifuged alternatively during 15 min, and finally trypsin was removed with magnetic separation. The supernatants were analyzed by MS spectrometry.

2.8. Mass spectrometry analysis

Prior to MALDI-TOF-MS analysis, the samples were mixed with 6 mg/ml of α -cyano-4-hydroxycinnamic acid matrix in 0.1% (v/v) TFA and 50% (v/v) ACN and 1 μl of each sample was hand-spotted in a MALDI target plate (384-spot Teflon[®]-coated plate) and allowed to air dry. Also, the samples digested with the two protocols were mixed in a ratio of 1:1 and were mixed with the matrix as described above. The samples were analyzed using a MALDI-TOF/TOF mass spectrometer model 4800 Proteomics Analyzer from Applied Biosystems (Framingham, MA) equipped with a Nd:YAG laser ($\lambda=355\text{ nm}$) and 4000 Series Explorer[™] software (Applied Biosystems). Internal calibration, data acquisition, processing and interpretation were carried out as recommended by the manufacturer. All mass spectra were externally calibrated using a standard peptide

mixture (Applied Biosystems). Mass spectral analysis for each sample was based on an average of 1500 laser shots.

2.9. Quantitative analysis (MS data processing)

Quantification was done using only peptides with double oxygen incorporation higher than 90%. All peptides were manually verified. The ratio was calculated as follows:

$$\text{Ratio} = I^{18}\text{O}/I^{16}\text{O}$$

where $I^{18}\text{O}$ corresponds to the intensity of the signal corresponding to the double-labelled peak whilst $I^{16}\text{O}$ corresponds to the intensity of the signal corresponding to the non-labelled peak.

3. Results and discussion

The conditions used to deliver ultrasonic energy into the samples to speed the digestion process were chosen based on our expertise with ultrasonic devices and their use in sample treatment for proteomics [3–7,12]. The ultrasonic amplitude was set at 50%, as it was demonstrated that this percentage of amplitude boosts protein digestion using direct high intensity ultrasonic energy, DHI-UE, when both, protein and enzyme are in the solution [3–7] or when the enzyme is immobilized in a solid support [10,12]. Nevertheless, for the last case, we have found recently that if the DHI-UE is replaced by the indirect high

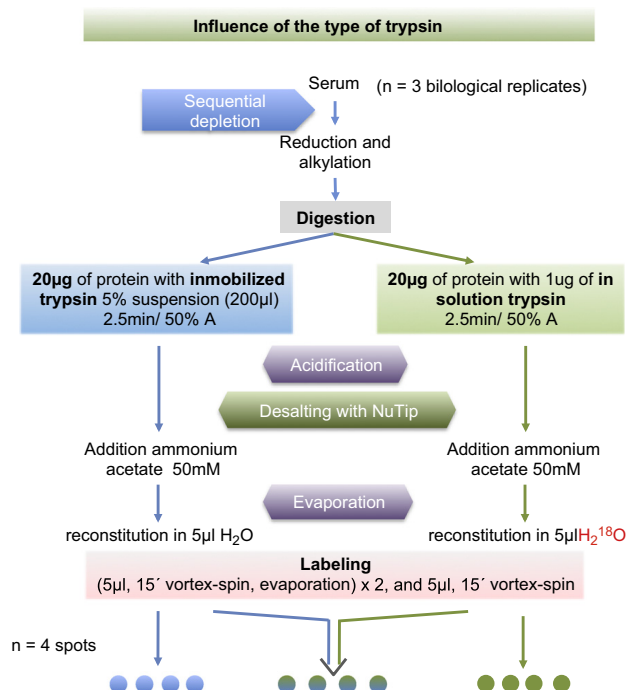


Fig. 1. Comprehensive scheme of the sample treatment followed to assess the influence of the type on trypsin used in a digestion done with the aid of an ultrasonic field provided by indirect high intensity ultrasonication (cup-horn device).

intensity ultrasonic energy, IHI-UE, the efficiency in the digestion is lowered dramatically. To demonstrate such findings we have followed the sample treatment outlined in Fig. 1. In brief, the sample treatment was done boosting the tryptic digestion with an ultrasonic cup-horn. The ultrasonic cup-horn provides an indirect high intensity ultrasonic energy, digesting the proteins (i) with in-solution trypsin or (ii) with trypsin immobilized on a solid support. To this end, the same sample was split into two subsamples. One subsample was digested with trypsin in solution and labelled with ^{18}O as explained below, whilst the other one was digested with immobilized trypsin and reconstituted in normal water. The aim was to compare the performance of both types of trypsin under the effects of IHI-UE.

3.1. Efficiency of the double ^{18}O incorporation

The efficiency in the digestion process was quantified using ^{18}O labelling. When the digestion of a proteome is done in the presence of ^{18}O -water, one or two oxygen atoms are incorporated into the peptide C-terminus. To ensure a correct quantitation, the sample treatment must enable double ^{18}O incorporation yielding higher than 90%, where otherwise quantification is compromised [3]. This is because isotopic peak overlapping may occur when the labelled and non-labelled samples are mixed, if double ^{18}O labelling has not been completed. This problem is described in Figure 1 of supplementary material, Fig. 1SM. In our approach, protein quantitation was done using three different serum samples. Each sample was split in two subsamples.

One was digested with immobilized trypsin in normal water, dried and then reconstituted in normal water. The other subsample was digested with in-solution trypsin in normal water, dried and then reconstituted in ^{18}O water. Then, peptides obtained for the first sub-sample were matched against the peptides obtained for the second one. Only common peptides with double ^{18}O incorporation higher than 90% were used to compare the performance of the type of trypsin as explained above. Results of this set of experiments are shown in Fig. 2. Through this approach the number of common peptides that were single or double labelled were 83 ± 11 , 44 ± 3 and 56 ± 10 ($n=4$, all peptides were manually verified, data presented in Fig. 3SM). From those peptides, the ones double labelled with efficiency higher than 90% were 60 ± 7 , 38 ± 3 and 49 ± 9 . Those peptides were then searched in 1:1 mixture to assess the influence of type of trypsin in the digestion efficiency as explained below.

3.2. Influence of the type of trypsin in the digestion efficiency

The efficiency of the digestion as a function of the trypsin used was assessed by searching in the 1:1 mixture the common peptides found with a degree of double incorporation higher than 90% as explained above. In-solution trypsin to immobilized trypsin ratios, were calculated as described in Section 2.9. Data presented in Fig. 3 show that for almost all these peptides this ratio was greater than 1.2, thus indicating that the serum treated with trypsin in solution was digested in more extent. For sample 1 (Fig. 3) from 23 peptides assessed, 22 presented a higher yield of digestion when in-solution

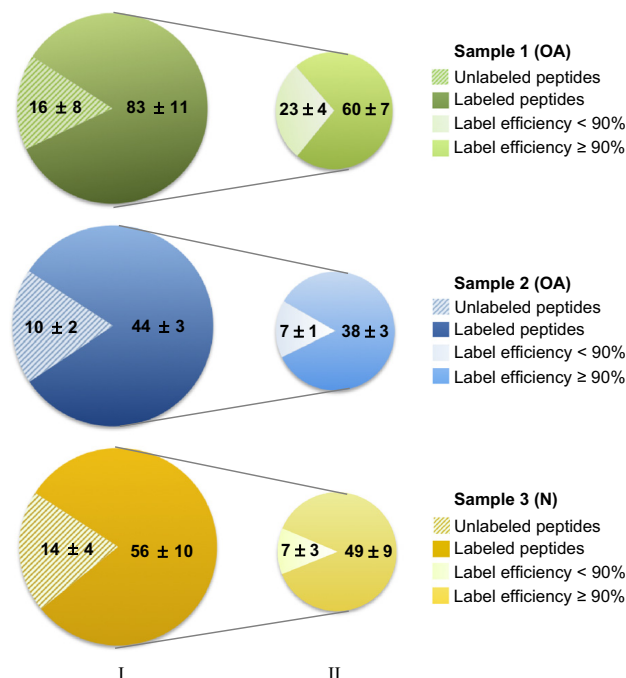


Fig. 2. I: Number of common peptides found in the 1:1 mixtures done mixing ^{18}O -label samples and their non label counterparts for three independent serum samples ($n=4$). II: For the same samples, number of peptides obtained with an ^{18}O double incorporation efficiency higher and lower than 90%.

trypsin was used. The digestion ratio was from 1.3 to 14.54 times higher. The same trend was observed for the other two samples. Thus, for sample 2, from 16 peptides assessed, 11 presented higher ratios of digestion, which varied from 1.2 to 4.74. For sample 3, from 20 peptides assessed 13 presented higher ratios, ranging from 1.3 to 8.37. It is noteworthy that a remarkable number of peptides found common between the ^{16}O and the ^{18}O digestion, were only found with the double ^{18}O incorporation in 1:1 mixture. This means that for these peptides the digestion with the in-solution trypsin was so efficient as compared to the digestion with immobilized trypsin that quantification was not possible, as the intensity corresponding to the non-labelled peptides was negligible. The number of these type of peptides was 29 ± 3 , 2 ± 2 and 15 ± 2 , corresponding to the three samples assessed.

3.3. Influence of the type of ultrasonic device.

As for the set of experiments explained above it was clear established that when IHI-UE is used as a tool to speed enzymatic

digestion, immobilized trypsin should not be used. We hypothesise that this finding is due to the aggregates formed by the beds were the trypsin was immobilized. Such aggregates are formed when an IHI-UE (cup-horn) is used to deliver ultrasound (see video 1SM). For instance, when an ultrasonic probe is used, such aggregates are not observed (see video 2SM). To further assess the influence of the type of tool used to deliver ultrasonic energy (direct or indirect) on the digestion process using immobilized trypsin, the ultrasonic probe (DHI-UE) and the cup-horn (IHI-UE) were compared as depicted in Fig. 2SM. The number of peptides chosen were 10 with labelling efficiencies higher than 90% (manually verified), which were used to calculate the ultrasonic probe to ultrasonic cup-horn ratio of digestion efficiency, in the same manner as explained in Section 2.9. Those 10 peptides were searched through four independent samples, which were treated by duplicate, being each duplicate spotted three times in the MALDI plate. This gives a total of 24 spectra. For the total of peptides assessed in these 24 spectra (240), the digestion efficiency was higher when an ultrasonic probe was used in a 70% of

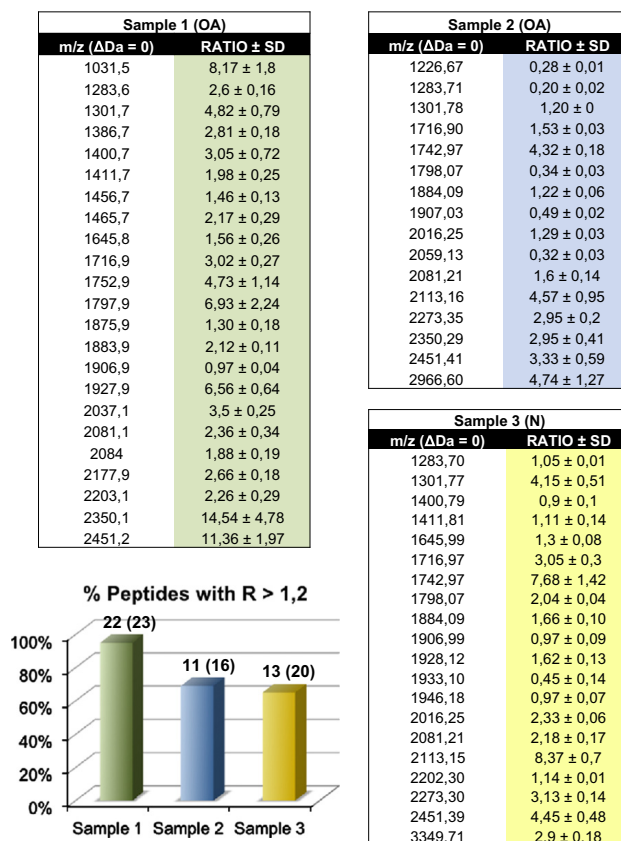


Fig. 3. The samples labelled and their non-labelled counterparts were compared to find common peptides following the sample treatment outlined in Fig. 1. Those peptides were then manually assessed in the label to no-label 1:1 mixtures to find out the label to no-label ratios. Data correspond to three different serum samples. Each sample was replicated four times.

the cases, with a ratio ranging from 1.2 to 4.4. A 15% of the peptides presented a ratio higher for the non-labelled ones, ranging from 1.25 to 2, whilst another 15% presented a 1:1 ratio.

4. Conclusions

It has been demonstrated that when trypsin immobilized on a solid support is used to digest proteomes under the effects of an ultrasonic field provided by high intensity indirect ultrasonic energy the performance is well below the one obtained when trypsin in solution is used. For immobilized trypsin higher digestion efficiencies are obtained if the process is aided with direct high intensity ultrasonic energy, e.g. ultrasonic probe, instead of doing it with indirect high intensity ultrasonic energy, e.g. ultrasonic cup-horn. As for the results reported in this work, quantification proteomics workflows aided with ultrasonic energy should be done with trypsin in solution. If immobilized trypsin is used, the assistance to the digestion process with ultrasonic energy should be done with a direct high intensity ultrasonic energy.

Acknowledgements

C. F. acknowledges the Foundation of the Complexo Universitario de A Coruña for financial support. H.M. Santos, acknowledges the post-doctoral grant SRFH/BPD/73997/2010, provided by Fundação para a Ciência e a Tecnologia—Ministério da Educação e Ciência, FCT-MEC, Portugal. J.L. Capelo and C. Lodeiro are grateful to the Xunta de Galicia (Spain) for the program Isidro Parga Pondal and for financial support provided under projects 09CSA043383PR and 10CSA383009PR-2010. C. Ruiz-Romero is supported by the Miguel Servet program from ISCIII-Fondo Investigación Sanitaria-Spain (CP09/00114). This study was supported by grants from

ISCIII-Fondo Investigación Sanitaria-Spain (CIBER-CB06/01/0040; RIER RD12/0009/0018, PI11/2397 and PI12/00329).

Appendix A. Supplementary information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.talanta.2012.12.022>.

References

- [1] D.B. Sprinson, D. Rittenberg, *Nature* 167 (1951) 484.
- [2] I.I. Stewart, T. Thomson, D. Figeys, *Rapid Commun. Mass Spectrom.* 15 (2001) 2456–2465.
- [3] J.L. Capelo, R.J. Carreira, L. Fernandes, C. Lodeiro, H.M. Santos, J. Simal-Gandara, *Talanta* 80 (2010) 1476–1486.
- [4] H.M. Santos, D. Glez-Peña, M. Reboiro-Jato, F. Fdez-Riverola, M. Diniz, C. Lodeiro, J.L. Capelo-Martínez, *Electrophoresis* 31 (2010) 3407–3419.
- [5] R.J. Carreira, C. Lodeiro, M. Reboiro-Jato, D. Glez-Peña, F. Fdez-Riverola, J.L. Capelo, *Talanta* 82 (2010) 587–593.
- [6] R.J. Carreira, C. Lodeiro, M.S. Diniz, I. Moura, J.L. Capelo, *Proteomics* 9 (2009) 49745159.
- [7] H.M. Santos, M. Reboiro-Jato, D. Glez-Peña, J.D. Nunes-Miranda, F. Fdez-Riverola, R. Carvallo, J.L. Capelo, *Talanta* 82 (2010) 1412–1420.
- [8] G. Wang, W.W. Wu, T. Pisitkun, J.D. Hoffert, M.A. Knepper, R.F. Shen, *Anal. Chem.* 78 (2006) 5752.
- [9] C. Wan, R.L. Cerny, D.S. Hage, *Anal. Chem.* 78 (2006) 7967.
- [10] G. Vale, H.M. Santos, R.J. Carreira, L. Fonseca, M. Miro, V. Cerda, M. Reboiro-Jato, J.L. Capelo, *Proteomics* 11 (2011) 3866–3876.
- [11] D. Lopez-Ferrer, J.L. Capelo, J. Vazquez, *J. Proteome Res.* 4 (2005) 1569–1574.
- [12] J.L. Capelo, R. Carreira, M. Diniz, L. Fernandes, et al., *Anal. Chim. Acta* 650 (2009) 151–159.
- [13] C. Fernandez, V. Calamia, P. Fernández-Puente, J.L. Capelo, C. Ruiz-Romero, F.J. Blanco, *Proteome Sci.* 10 (2012) 55.
- [14] S.E. Warder, L.A. Tucker, T.J. Streltzer, E.M. McKeegan, J.L. Meuth, P.M. Jung, A. Saraf, B. Singh, J. Lai-Zhang, G. Gagne, J.C. Rogers, *Anal. Biochem.* 387 (2009) 184–193.
- [15] R. Kay, C. Barton, L. Ratcliffe, B. Matharoo-Ball, P. Brown, J. Roberts, P. Teale, C. Creaser, *Rapid Commun. Mass Spectrom.* 22 (2008) 3255–3260.

3.4. Capítulo IV: Depleción secuencial acoplada a extracción secuencial en C18 como herramienta rápida para la búsqueda de biomarcadores de enfermedades reumáticas en suero humano

La eliminación de las proteínas mayoritarias del suero se debe realizar escogiendo el método de depleción más adecuado para el perfil proteico de la enfermedad a estudiar. Bajo esta premisa, se han desarrollado y utilizado en proteómica numerosos métodos de depleción química. La depleción con acetonitrilo (ACN) proporciona un suero rico en apolipoproteínas, que resulta adecuado para estudios relacionados con enfermedades cardiovasculares. En cambio, la depleción química con ditioneitol (DTT) enriquece el suero en inmunoglobulinas, apropiado para estudios con mielomas y linfomas. Además, el método con DTT parece ser el que mejor depleciona las proteínas abundantes a la vez que preserva un gran número de proteínas en solución. Recientemente se ha aplicado con éxito, en nuestro laboratorio, la combinación de la depleción con ACN y con DTT de forma secuencial en la búsqueda de biomarcadores de artrosis (OA).

El empleo de métodos de espectrometría de masas basados en el perfil peptídico (también llamados de huella peptídica) se ha incrementado en los diez últimos años por ser métodos rápidos para el estudio de los cambios en los patrones de expresión proteica de sistemas biológicos, ayudando a la detección temprana de enfermedades, a diferenciar sus estados, a monitorizar terapias y a la prognosis de enfermedades malignas. Antes de la espectrometría de masas se suele realizar un paso previo con el fin de desalar y preconcentrar los péptidos mediante Zip-Tips (extracción en fase sólida) o dispositivos similares (tipo NuTips). Adicionalmente, los dispositivos de tipo Zip-Tip también pueden utilizarse para fraccionar la muestra y reducir así su complejidad: en este caso los péptidos unidos a estas resinas se eluyen de forma secuencial empleando disoluciones de diferente composición, de modo que cada disolución extraerá diferentes grupos de péptidos retenidos en la resina, que podrán estudiarse de forma independiente.

El objetivo de este trabajo fue acoplar, por primera vez, los pasos anteriormente mencionados para desarrollar un método sencillo, rápido y económico de búsqueda de posibles biomarcadores proteómicos en muestras de suero humano. Para ello, se propuso realizar una depleción química, seguida de una digestión y

posterior separación de los péptidos utilizando NuTips mediante eluciones secuenciales con diferentes concentraciones de ACN, para finalmente obtener los perfiles de masas de cada eluido mediante espectrometría de masas tipo MALDI-TOF/TOF.

Se realizó un primer experimento en donde se deplecionaron sueros con DTT y se digirieron con la tripsina en solución acelerada por ultrasonidos. Dos cantidades distintas de proteína digerida (péptidos) se cargaron en NuTips (con resina C18) y se eluyeron con 10 soluciones de ACN de % creciente, con el fin de comprobar la capacidad de carga de dicha resina. Teniendo en cuenta el total de proteínas identificadas, en ambos casos se llegó a la conclusión de que no se debe sobrepasar la capacidad de carga para no perder información. También se observó que el número de proteínas identificadas aumentó hasta el 35 % de ACN, de modo que las eluciones con % de ACN mayor no eran necesarias.

De acuerdo con los anteriores resultados, se realizó un segundo experimento con el mismo tipo de depleción (DTT) y digestión, pero en este caso se escogió la cantidad de proteína digerida para cargar en las NuTips y también se modificó el gradiente de ACN para aumentar la información obtenida en el espectrómetro. El número total de proteínas identificadas resultó el mismo que en el anterior experimento (21). En estos dos primeros análisis, el 57 % de las proteínas identificadas resultaron ser de las más abundantes del suero.

Para evitar la presencia tan elevada de proteínas abundantes en los perfiles proteicos obtenidos, se realizó otro experimento en donde se sustituyó la depleción con DTT por la depleción secuencial (DTT + ACN) previamente descrita. Se identificó un número de proteínas similar (22) pero sólo un 23 % de ellas pertenecían al grupo de las abundantes. También se estudiaron los % de ACN para seleccionar los que mas información aportaron.

Con todos estos resultados se escogió el método de depleción secuencial y un gradiente de ACN más corto (0; 4; 7; 10; y 14 %) para crear un método final de búsqueda de biomarcadores.

Para comprobar la aplicación del método a la búsqueda de biomarcadores, se realizaron dos experimentos distintos en sueros de pacientes con artrosis (OA) y de donantes control (N). El primero de ellos se realizó en 10 sueros agrupados en 2 muestras (5 sueros/muestra) de cada condición, y se obtuvieron los perfiles proteicos de cada fracción eluída. Se observó un patrón de proteínas distinto en cada % de elución para cada grupo de donantes.

De acuerdo con el anterior resultado, se realizó un último análisis para obtener el perfil peptídico de cada fracción en lugar del perfil proteico. Para ello, se seleccionaron 20 sueros de donantes OA, 20 sueros de donantes control, 20 sueros de donantes con artritis psoriásica (APS) y 20 sueros de donantes con artritis reumatoide (AR). Las muestras se procesaron individualmente, se deplecionaron, se realizó la digestión y después la separación con el gradiente de ACN escogido. Cada eluído se analizó por espectrometría de masas por quintuplicado y se obtuvieron los perfiles peptídicos de cada condición mediante herramientas informáticas. Finalmente, se escogieron los péptidos capaces de discriminar entre las distintas condiciones con significancia estadística.

Los resultados obtenidos demostraron que el método desarrollado fue útil para encontrar diferencias entre los cuatro grupos de muestras, permitiendo la clasificación de las mismas mediante los péptidos seleccionados. Además, cabe destacar que el material empleado es económico, y además se trata de un método sencillo y rápido que se puede implementar fácilmente en cualquier estudio proteómico de otras patologías o condiciones en muestras de suero humano.

3.4.1 Materiales y métodos

Muestras de suero

Las muestras de suero humano se obtuvieron de donantes anónimos en el Complejo Hospitalario Universitario de A Coruña, España. Todos los pacientes firmaron el consentimiento informado. El estudio fue aprobado por el comité de ética local en Galicia (España). Se seleccionaron un total 80 muestras de suero, 20 muestras de cada uno de los siguientes grupos: donantes con artrosis (OA), donantes con artritis reumatoide (AR), donantes con artritis psoriásica (APS) y controles sanos (N).

Las muestras de suero se almacenaron a -80 ° C hasta su procesamiento.

Depleción con DTT

La depleción con DTT se realizó de acuerdo con el protocolo descrito por Warder y colaboradores [121]. Brevemente, se añadieron 2.2 μ l de DTT 500 mM a 20 μ l de suero y se agitó en vórtex. La muestra se incubó 1 h hasta la aparición de un precipitado blanco viscoso, a continuación se centrifugó a 14000 g durante 20 min dos veces. El sobrenadante se transfirió a un tubo limpio LoBind (Eppendorf) y se llevó a sequedad en una centrifuga de vacío (Savant SPD121P Speed Vac de Thermo, Waltham, EE.UU.).

Depleción secuencial de proteínas en suero

La depleción de proteínas se realizó de acuerdo con el protocolo desarrollado previamente en nuestro laboratorio (Artículo II). Cada uno de los sueros fueron sometidos a un protocolo de depleción secuencial que implica dos etapas de precipitación: primero con DTT y a continuación con ACN, de acuerdo con el protocolo descrito por Kay y colaboradores [122], con pequeñas modificaciones. 20 μ l de suero se deplecionaron con DTT y el sobrenadante se deplecionó de nuevo mediante la adición de 57 % (v/v) acetonitrilo, seguido de dos ciclos de vórtex breve y 10 min de ultrasonidos en un baño de ultrasonidos (SONOREX, Bandelin). La muestra se centrifugó a 14000 g durante 10 min, y el sobrenadante se transfirió a un tubo limpio LoBind y se evaporó a sequedad empleando una centrifuga de vacío.

Digestión en solución

La digestión en solución acelerada con ultrasonidos se realizó de acuerdo con el protocolo de digestión proteolítica ultrarrápida desarrollado previamente [123]. La muestra evaporada se resuspendió en 20 μ l de bicarbonato amónico 25 mM (AmBi), se añadieron 10 μ l de acetonitrilo y se mezcló mediante vortex y sonicación durante 1 min (50 % de amplitud) en baño de ultrasonidos. Los residuos de cisteína de las proteínas se redujeron con 2 μ l de DTT 110 mM seguido de vortex y 1 min de sonicación, y a continuación se bloqueó con 2 μ l de IAA 600 mM, vortex y 1 min sonicación. Para inactivar la IAA se añadieron a la muestra 10 μ l de DTT 110 mM. Después, la muestra se diluyó hasta un volumen final de 200 μ l con AmBi 12.5 mM, y se añadió tripsina de acuerdo con un ratio 1:20 (p/p) de tripsina:proteína. La digestión se llevó a cabo en el aparato de ultrasonidos indirectos de alta intensidad (SONOPULS HD 2200 con accesorio BB6, Bandelin) durante 5 min a un 50 % de amplitud. Finalmente, se añadieron 2 μ l de ácido fórmico al 50 % (v/v) para parar la reacción enzimática. El suero digerido se evaporó hasta sequedad en una centrifuga de vacío.

Elución secuencial mediante NuTips (separación de péptidos)

Las muestras digeridas se reconstituyeron en 30 μ l de TFA al 0.1 % (v/v), y diferentes cantidades de proteína (30, 20 y 18 μ g) se cargaron en una punta con octadecilo (C18) inmovilizado de tipo NuTip C18 large (Glygen, Columbia, EE.UU.). La separación de los péptidos se realizó de acuerdo con el siguiente protocolo:

- Acondicionado: se aspiró y expulsó 5 veces 30 μ l de una solución de ACN al 60-80 % (v/v) y TFA al 0.1 % (v/v), y se lavó 3 veces con TFA al 0.1 %.
- Unión: se aspiró y expulsó la muestra 50 veces, para permitir que los péptidos se adsorbieran al material de fase reversa empaquetado en la punta.
- Elución secuencial: se aspiró y expulsó (20 veces cada uno) 30 μ l de diferentes porcentajes de ACN, de baja a alta concentración de ACN. Los eluidos se evaporaron a sequedad hasta el análisis mediante espectrometría de masas.

Siguiendo este protocolo, se realizaron los siguientes experimentos para optimizar las condiciones de la elución secuencial:

- 1) Primer experimento: un conjunto (pool) de 5 sueros de donantes artrósicos se deplecionó con DTT y se digirió. A continuación, 30 μ g de proteína digerida se cargaron por quintuplicado en una NuTip C18 large y la elución se realizó con 0,

10, 15, 20, 25, 30, 35, 40, 55 y 80 % (v/v) de ACN. Cada eluido se recogió en un tubo nuevo y se llevó a sequedad en una centrifuga de vacío. El experimento se repitió por triplicado con 18 µg de proteína digerida.

2) Segundo experimento: un conjunto de 5 sueros de donantes artrósicos se deplecionó con DTT y se digirió. 20 µg de proteína digerida se cargaron por triplicado en una NuTip C18 large y la elución se realizó con un nuevo gradiente de 0, 4, 7, 10, 14, 17, 20, 25, 35 y 60 % (v/v) de ACN. Cada eluido se recogió en un tubo nuevo y se llevó a sequedad en una centrifuga de vacío.

3) Tercer experimento: un conjunto de 5 sueros de donantes artrósicos se deplecionó con el protocolo de la depleción secuencial. A continuación, 20 µg de proteína digerida se cargaron por triplicado en una NuTip C18 large y la elución se realizó con un nuevo gradiente de ACN: 0, 4, 7, 10, 14, 17, 20, 25, 35 y 60 % (v/v). Cada eluido se recogió en un tubo nuevo y se llevó a sequedad en una centrifuga de vacío.

Comparación de los perfiles proteicos de sueros de donantes control y de pacientes con artrosis

Se agruparon volúmenes iguales de sueros de 10 pacientes artrósicos en dos conjuntos de 5 muestra cada uno, y se agruparon de la misma forma 10 sueros de donantes controles. Las muestras agrupadas se deplecionaron con el protocolo de depleción secuencial y se digirieron. A continuación, 20 µg de proteína digerida se cargaron en una NuTip C18 large y la elución se realizó con 0, 4, 7, 10, 14 y 35 % (v/v) de ACN. Cada eluido se recogió en un tubo nuevo y se llevó a sequedad en una centrifuga de vacío hasta el análisis mediante espectrometría de masas.

Obtención de los perfiles peptídicos de muestras de sueros OA, APS, AR y N

20 sueros de cada condición, OA, AR, APS y N, se trataron individualmente con la depleción secuencial y la digestión en solución. A continuación, 20 µg de proteína digerida se cargaron en una NuTip C18 large y la elución se realizó con 4, 7, 10 y 14 % (v/v) de ACN. Los eluidos se recogieron en un tubo nuevo y se llevaron a sequedad en una centrifuga de vacío. Por último, las muestras se resuspendieron en 6 µl de TFA 0.1 % (v/v) y se depositó 1 µl de cada muestra en la placa MALDI por quintuplicado para el análisis en el espectrómetro de masas.

Análisis mediante espectrometría de masas

Las muestras evaporadas se resuspendieron con ácido trifluoroacético (TFA) 0.1% (v/v) y 1 µl de cada muestra se depositó en la placa MALDI (384 puntos recubierta de Teflón®), se dejaron secar al aire a temperatura ambiente. Posteriormente, se añadieron a las muestras secas 1 µl de una solución de 3 mg/ml de la matriz de ácido α -ciano-4-hidroxicinámico en TFA al 0.1% (v/v) y 50% (v/v) ACN, y se dejaron secar de nuevo.

Las muestras se analizaron en modo de ion positivo usando un espectrómetro de masas MALDI-TOF/TOF 4800 Proteomics Analyzer (ABSciex, Framingham, MA, EE.UU.), equipado con un láser Nd:YAG ($\lambda=355$ nm) y el programa 4000 Series ExplorerTM (ABSciex). La calibración interna, adquisición de datos, procesamiento e interpretación se llevaron a cabo según lo recomendado por el fabricante. Todos los espectros de masas deben ser calibrados externamente empleando una mezcla de péptidos estándar (ABSciex). El análisis del espectro de masas para cada muestra corresponde a un promedio de 1.500 disparos de láser.

Los espectros de fragmentación para la identificación de las proteínas se adquirieron seleccionando los 20 iones más abundantes de cada mapa de masas MALDI-TOF y con un promedio de 2000 disparos de láser por espectro. Todos los picos con ratio señal/ruido (S/N) superior a 10 se incluyeron en la búsqueda en bases de datos. La identificación de las proteínas se realizó empleando el motor de búsqueda MASCOT (Matrix Science) con los siguientes parámetros: (i) Base de datos: Swiss-Prot 2011 (532792 secuencias, 188961396 residuos), (ii) taxonomía: Homo sapiens (20257 secuencias), (iii) enzima: tripsina, (iv) un error en la acción del enzima, (v) modificaciones fijas: carbamidometilación en cisteínas (C), (vi) modificaciones variables: oxidación en metionina (M), (vii) tolerancia en la masa del péptido: 50 ppm, (viii) tolerancia en la masa del fragmento: 0.3 Da. Sólo se consideraron las identificaciones con $p < 0.05$.

Los espectros de fragmentación para la secuenciación de los péptidos seleccionados se adquirieron mediante la selección de los precursores de interés en cada mapa de masas MALDI-TOF, con un promedio de 2000 disparos de láser por espectro. Para la fragmentación (MS/MS), todos los picos con un ratio señal/ruido (S/N) > 10 se incluyeron en la búsqueda en bases de datos. La secuenciación de péptidos se realizó mediante el software ProteinPilotTM (ABSciex), con los siguientes parámetros: (i) versión UniProt: 2013_09, que contiene 540.958 entradas de secuencia, (ii) taxonomía: Homo sapiens, (iii)

enzima: tripsina o sin enzima, (iv) modificaciones fijas: iodoacetamida. Sólo se informó de los péptidos identificados con al menos un 70 % de confianza.

Análisis de los datos de los sueros individuales de donantes con OA, AR, APS y N

Cada espectro se pre-procesó empleando los siguientes parámetros: rango de masas de 500 a 4000 Da, densidad máxima de picos: 10 picos por 200 Da, relación mínima de señal/ruido (S/N): 10, área mínima de 100, número máximo de picos por espectro: 120. Los picos se alinearon con una tolerancia en la masa del péptido de 150 ppm, y finalmente se creó un espectro representativo para cada muestra con aquellos picos presentes en al menos 4 de las 5 réplicas espectrales de la muestra. Los picos resultantes fueron filtrados mediante un test de Fisher corregido con la FDR de Benjamini Hochberg, seleccionando aquellos con un valor $q < 0.05$.

Se realizó un análisis de componentes principales (PCA) sobre los espectros de masas de los sueros de donantes control y artrósicos. Antes de realizar el PCA, se creó un grupo de datos único para cada porcentaje de elución (4, 7, 10 y 14 %), con el espectro representativo de cada muestra. Los espectros de cada grupo de datos se convirtieron en vectores 1s y 0s, donde 1 significa presencia de pico y 0 significa ausencia de pico. El análisis PCA y su visualización se realizaron usando el programa RapidMiner v5.3 (<http://rapid-i.com/content/view/181/190/>), que configura el algoritmo PCA para reducir la dimensionalidad de los grupos de datos hasta 3 componentes principales.

Para la búsqueda de biomarcadores de sueros de las distintas enfermedades, se seleccionaron los péptidos capaces de diferenciar las cuatro condiciones a partir de los espectros representativos que se crearon. Primero se seleccionaron los péptidos que se detectan con una mayor frecuencia en una condición con respecto a las otras de forma significativa, de acuerdo con un $p < 0.05$ del test de Fisher aplicado por parejas. A continuación se seleccionaron los péptidos que se detectan con mayor frecuencia en más de dos condiciones con respecto a las otras, cumpliendo el mismo criterio de p valor inferior a 0.05.

3.4.2. Resultados

Separación de los péptidos: depleción con DTT

El tratamiento de las muestras se representa en el esquema de la figura 1. Las muestras se deplecionaron con DTT, se digirieron en solución y los péptidos se separaron en las puntas con C18 inmovilizado empleando el gradiente de ACN descrito.

Se cargaron dos cantidades diferentes de proteína digerida en las NuTips para comprobar su capacidad de carga máxima, 30 y 18 μg . Con 30 μg se identificaron un total de 14 proteínas distintas que se muestran en la figura 2 (13 ± 1 , rango entre 12-14, $n = 5$), de las cuales 9 pertenecían al grupo de las 20 proteínas más abundantes (top 20). Cuando se cargaron 18 μg de proteína en las NuTips, se incrementó el número de proteínas distintas identificadas a 21 (19 ± 2 , rango entre 16-20, $n = 3$), 12 de ellas pertenecían al grupo de las top 20. El total de las proteínas distintas identificadas no aumentó a % de ACN mayores del 35 % (v/v) (figura 3). Con el nuevo gradiente de elución (comprendido entre 0 y 60 %) se cargaron 20 μg de proteína en las NuTips y se identificaron 21 proteínas (19 ± 1 , rango entre 18-20, $n = 3$), el 57 % (12) incluidas en el grupo de las top 20.

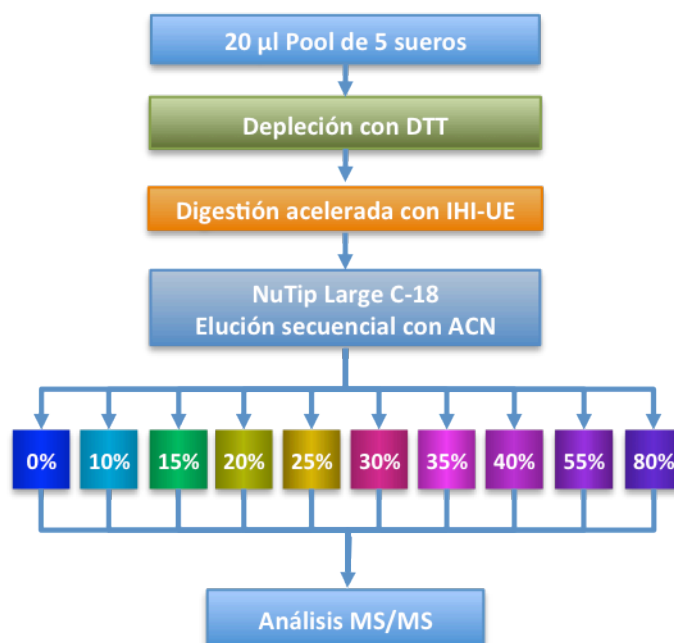


Figura 1. Esquema del tratamiento de las muestras correspondiente al primer experimento con la depleción con DTT y la elución secuencial. IHI-UE: ultrasonidos indirectos de alta intensidad.

En la figura 2 se representa la reproducibilidad del método de separación secuencial de los péptidos empleando NuTips. La reproducibilidad fue superior al 80 % en las cinco réplicas técnicas realizadas con los 30 μ g de proteína.

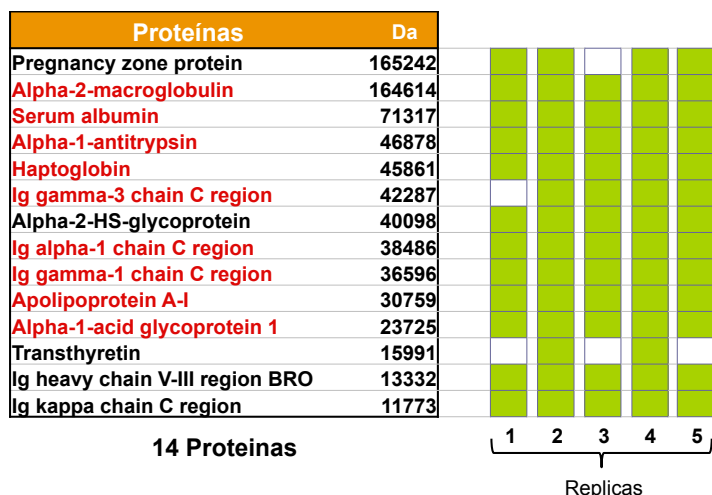


Figura 2. Reproducibilidad de la separación de los péptidos con la elución secuencial. Proteínas identificadas mediante sus espectros de fragmentación (MS/MS) en las 5 réplicas realizadas del mismo conjunto de 5 sueros de donantes artrósicos. Las proteínas en rojo pertenecen al grupo de las 20 más abundantes (top 20).

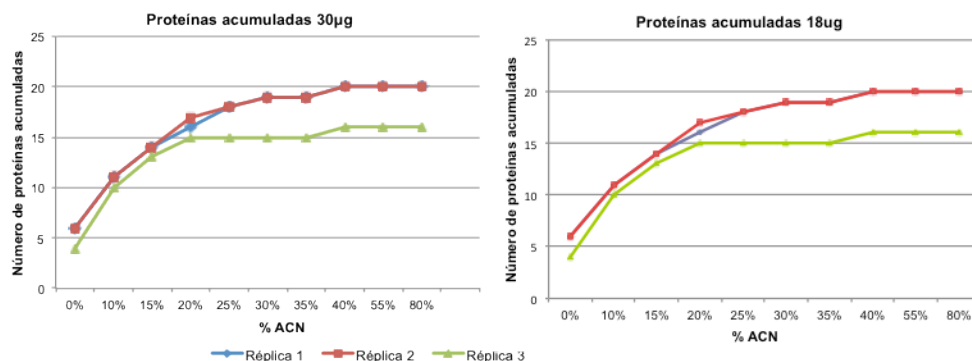


Figura 3. Número de proteínas acumuladas que se identificaron en cada porcentaje de elución con ACN, en los primeros experimentos en los que se deplecionaron las muestras con DTT.

Separación de los péptidos: depleción secuencial

A continuación, las muestras se deplecionaron con el protocolo de depleción secuencial, se digirieron en solución y 20 µg de proteína se separaron con el gradiente de ACN descrito. Con este método se identificaron 22 proteínas, de ellas sólo 5 pertenecían al grupo de las top 20. En la figura 4 se muestran las proteínas identificadas en las tres réplicas técnicas. Todas las proteínas se identificaron en el rango de concentración de ACN comprendido entre 0 y 35 % (v/v). El 82 % de las proteínas se identificó en un mínimo de dos de las tres réplicas.

En la figura 5 se muestran las proteínas identificadas con los dos métodos de depleción, 21 con la depleción con DTT y 22 con la depleción secuencial. Sólo 5 proteínas se identificaron con los dos métodos. En cuanto a las proteínas abundantes, se detectaron 12 en el método con DTT y 5 con la depleción secuencial. En la tabla 1 se encuentran las proteínas identificadas en cada % de elución para las tres réplicas.

Proteínas identificadas (n=3)				
A1AG1	Alpha-1-acid glycoprotein 1			
A1AG2	Alpha-1-acid glycoprotein 2			
A1AT	Alpha-1-antitrypsin			
FETUA	Alpha-2-HS-glycoprotein			
ANGT	Angiotensinogen			
APOA1	Apolipoprotein A-I			
APOA2	Apolipoprotein A-II			
APOA4	Apolipoprotein A-IV			
APOC3	Apolipoprotein C-III			
APOD	Apolipoprotein D			
XIAP	Baculoviral IAP repeat-containing protein 4			
CCD78	Coiled-coil domain-containing protein 78			
HBB	Hemoglobin subunit beta			
IGKC	Ig kappa chain C region			
KNG1	Kininogen-1			
A2GL	Leucine-rich alpha-2-glycoprotein			
KDM4D	Lysine-specific demethylase 4D			
MARHA	Probable E3 ubiquitin-protein ligase MARCH10			
THRB	Prothrombin			
RHXF1	Rhox homeobox family member 1			
SAA	Serum amyloid A protein			
TF3B	Transcription factor IIIB 90 kDa subunit			
22 Proteínas		1	2	3
		C18 réplicas		

Figura 4. Proteínas identificadas a partir de los espectros de fragmentación de las muestras deplecionadas con el protocolo de depleción secuencial. En rojo se distinguen las proteínas que pertenecen a las top 20 del suero.

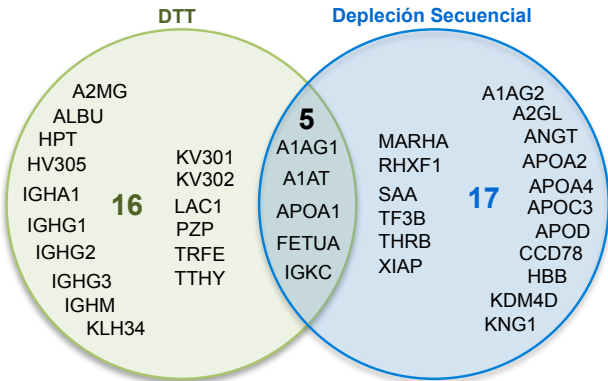


Figura 5. Los diagramas de Venn muestran las proteínas identificadas con los dos métodos de depleción empleados. Las abreviaturas se describen en la tabla 3.

C-18 Depleción secuencial (DTT + ACN)					
% ACN	Réplica 1		Réplica 2		Réplica 3
0%	APOA2	A1AG1	APOA1 A1AG1	APOA2 KNG1	
4%	APOA2 APOA1 A1AG1 A2GL APOD	KDM4D APOD TF3B KNG1 CCD78 APOC3	APOA2 APOA1 A1AG1 KNG1	APOA4 A2GL APOD	APOA2 APOA1 APOA4 A2GL RHXF1 A1AG1 APOD CCD78 KNG1
7%	APOA1 APOC3 A1AG1 APOA2 HBB	A2GL APOD CCD78 KNG1 MARHA	APOA1 APOA2 A1AG1	APOA4 A2GL XIAP	APOA2 APOA1 APOA4 A2GL APOD A1AG1 APOC3 RHXF1 MARHA CCD78
10%	APOA1 A1AG1 SAA IGKC	APOA2 A2GL FETUA	APOA1	A2GL	APOA1 A1AG1 FETUA A2GL
14%	FETUA A1AG2 A1AG1 THRB	APOD APOA1 HBB	A1AG1 A1AG2	APOA1 FETUA	A1AG1 FETUA A1AG2 APOD APOA1
17%	A1AG1 A1AG2 FETUA APOA1	APOD THRB HBB	A1AG1 A1AG2 APOA1	HBB FETUA APOD	A1AG1 A1AG2 APOD FETUA APOA1
20%	A1AG1 A1AG2 APOD	HBB FETUA	A1AG1		A1AG1 A1AG2 APOD FETUA
25%	A1AG1	IGKC	A1AG1 A1AT	IGKC	A1AG1 IGKC
35%	A1AG1 A1AT IGKC	SAA ANGT	A1AG1		A1AG1
60%	A1AG1	SAA			

Tabla 1. Proteínas identificadas mediante espectrometría de masas en tándem (MS/MS) en los diferentes % de elución y en las tres réplicas técnicas.

Comparación de los perfiles proteicos de sueros de donantes control y de pacientes con artrosis

Cada conjunto de muestras, 2 de sueros de donantes artrósicos y 2 de sueros de donantes sanos (controles) se trataron con la depleción secuencial, se digirieron y se separaron con un gradiente corto de ACN (figura 6). En la tabla 2 y en los diagramas de Venn que se muestran en la figura 7 se describen las proteínas identificadas en cada % de ACN para las dos condiciones. Los datos correspondientes a las identificaciones obtenidas con MASCOT se muestran en el material suplementario del capítulo (anexo IV).

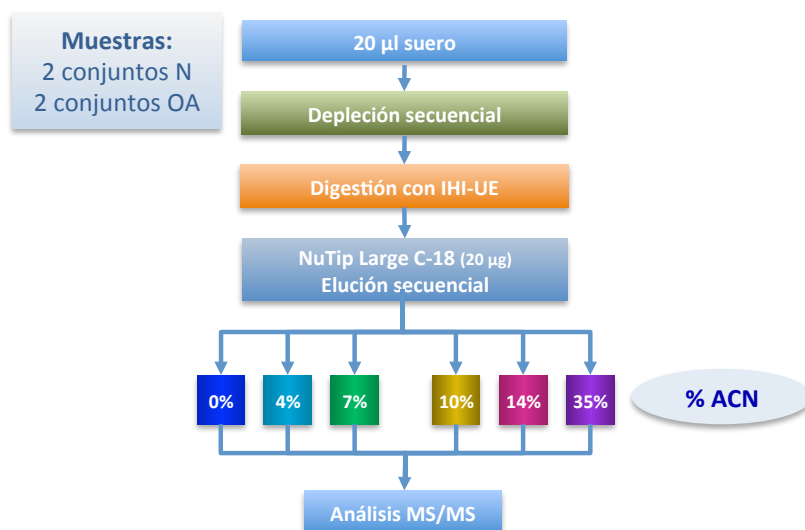


Figura 6. Esquema del flujo de trabajo seguido con las muestras de grupos de pacientes artrósicos (OA) y de donantes controles sanos (N), para la comparación de los perfiles proteicos obtenidos. IHI-UE: ultrasonidos indirectos de alta intensidad.

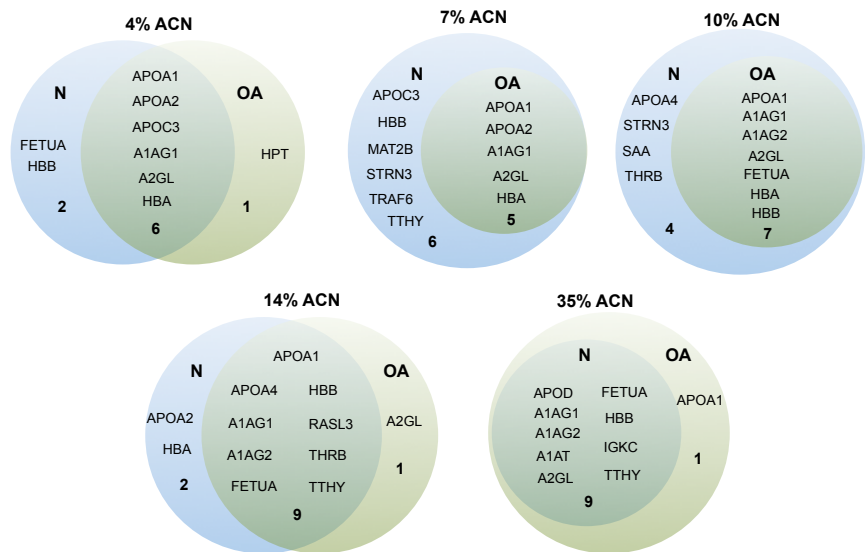


Figura 7. Diagramas de Venn con las proteínas identificadas en los pools de los sueros de donantes artrósicos (OA) y de controles (N). Las muestras se deplecionaron con el protocolo de depleción secuencial, se digirieron y se separaron los péptidos con la elución secuencial con diferentes % (v/v) de ACN.

% ACN	Normal 1		Normal 2		OA 1	OA 2	
0%			APOA2				
4%	APOA2 HBA HBB A1AG1	A2GL APOA1 APOC3	APOA1 A1AG1 APOC3	A2GL APOA2 FETUA	APOA2 HPT	APOA2 APOA1 A1AG1	APOC3 A2GL HBA
7%	APOA1 APOA2 APOC3 HBB	A1AG1 A2GL HBA TRAF6 STRN3	APOA1 APOC3 APOA2	TTHY A1AG1 A2GL TRAF6 MAT2B	APOA1 APOA2 A1AG1 HBA A2GL	APOA1 APOA2 A1AG1	HBA A2GL
10%	HBB APOC3 A1AG1 A1AG2 HBA	A2GL FETUA THRB STRN3	APOA4 A1AG1 A1AG2 FETUA	APOA1 A2GL THRB HBB SAA	HBA APOA1 A1AG1 A2GL FETUA	APOA1 A1AG2 HBB	HBA A2GL
14%	HBB FETUA A1AG1 A1AG2	HBA TTHY APOA4 THRB	APOA1 FETUA A1AG1 A1AG2 HBB	APOD APOA4 TTHY THRB RASL3	HBB A1AG1 APOA1 FETUA THRB A2GL	A1AG1 A1AG2 HBB	THRB APOA4 TTHY RASL3
35%	A1AG1 A1AG2 HBB TTHY	APOD A2GL IGKC A1AT	A1AG1 A1AG2 HBB TTHY	APOD A2GL IGKC FETUA	HBB A1AG1 APOA1 A2GL IGKC FETUA	A1AG1 A1AG2 HBB APOA1	TTHY APOD IGKC A1AT

Tabla 2. Proteínas identificadas mediante los espectros de fragmentación de los pools de muestras de donantes normales (controles) y de muestras de donantes artrósicos (OA).

Clasificación de las muestras de suero mediante su perfil peptídico

Con el fin de verificar la utilidad de este flujo de trabajo para la búsqueda de biomarcadores de enfermedades reumáticas, las 80 muestras de suero (20 de cada condición: OA, AR, APS y controles) fueron procesadas individualmente siguiendo esta estrategia: depleción secuencial, digestión en solución y separación de los péptidos mediante eluciones con cuatro porcentajes de ACN, y por último se analizaron por quintuplicado en el espectrómetro de masas (figura 8). Con los péptidos de los espectros de cada % de elución para las muestras de donantes sanos y artrósicos se realizó un análisis de componentes principales (PCA) en el que las muestras se separaron en las dos condiciones en al menos dos planos para los cuatro porcentajes de elución (figura 9).

De los espectros representativos que se obtuvieron para cada muestra se seleccionaron un total de 107 péptidos relevantes en una o en varias condiciones con una $p < 0.05$ del test de Fisher aplicado por parejas (tabla 3). Los péptidos se dividieron en los siguientes grupos: (i) 13 péptidos más frecuentes en los sueros de donantes artrósicos comparado con las otras tres condiciones, con un p valor significativo cuando se enfrentó la condición de artrosis frente a cada una de las otras condiciones; (ii) 13 péptidos más frecuentes en los sueros de donantes sanos (controles) comparado con las otras tres condiciones, con un p valor significativo cuando se enfrentó la condición de control frente a cada una de las otras condiciones; (iii) 12 péptidos más frecuentes en los sueros de donantes con artritis psoriásica comparado con las otras tres condiciones, con un p valor significativo cuando se enfrenta la condición de artritis psoriásica frente a cada una de las otras condiciones; (iv) 13 péptidos más frecuentes en los sueros de donantes con artritis reumatoide comparado con las otras tres condiciones, con un p valor significativo cuando se enfrenta la condición de artritis reumatoide frente a cada una de las otras condiciones; (v) 14 péptidos más frecuentes en los sueros de donantes con artritis psoriásica y en los sueros con artritis reumatoide comparado con las otras dos condiciones, con un p valor significativo cuando se enfrentan APS y AR con OA y N; (vi) 4 péptidos más frecuentes en los sueros de donantes artrósicos y de donantes normales comparado con las otras dos condiciones, con un p valor significativo cuando se enfrentan OA y N con APS y AR; (vii) 1 péptido más frecuente en los sueros de donantes con artrosis y en los sueros de donantes artritis reumatoide, con un p valor significativo cuando se enfrentan OA y AR con N y APS; (viii) 3 péptidos menos frecuentes en los sueros de donantes con artritis

psoriásica comparado con las otras tres condiciones, con un p valor significativo cuando se enfrenta la condición de artritis psoriásica frente a las otras tres condiciones; (ix) 34 péptidos menos frecuentes en los sueros de donantes control comparado con las otras tres condiciones, con un p valor significativo cuando se enfrenta la condición control frente a las otras tres condiciones.

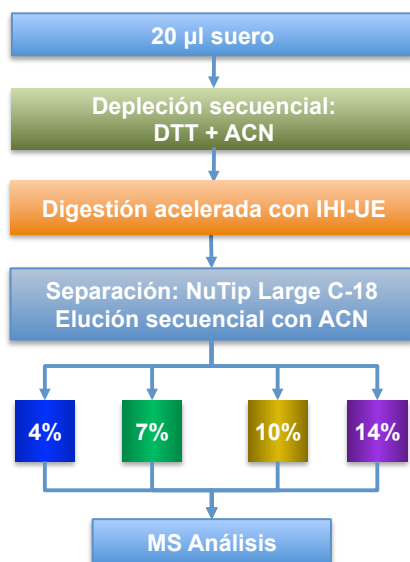


Figura 8. Esquema del flujo de trabajo seguido con las 80 muestras de los sueros de donantes artrósicos, con artritis psoriásica, con artritis reumatoide y controles, para la comparación de los perfiles peptídicos obtenidos. IHI-UE: ultrasonidos indirectos de alta intensidad.

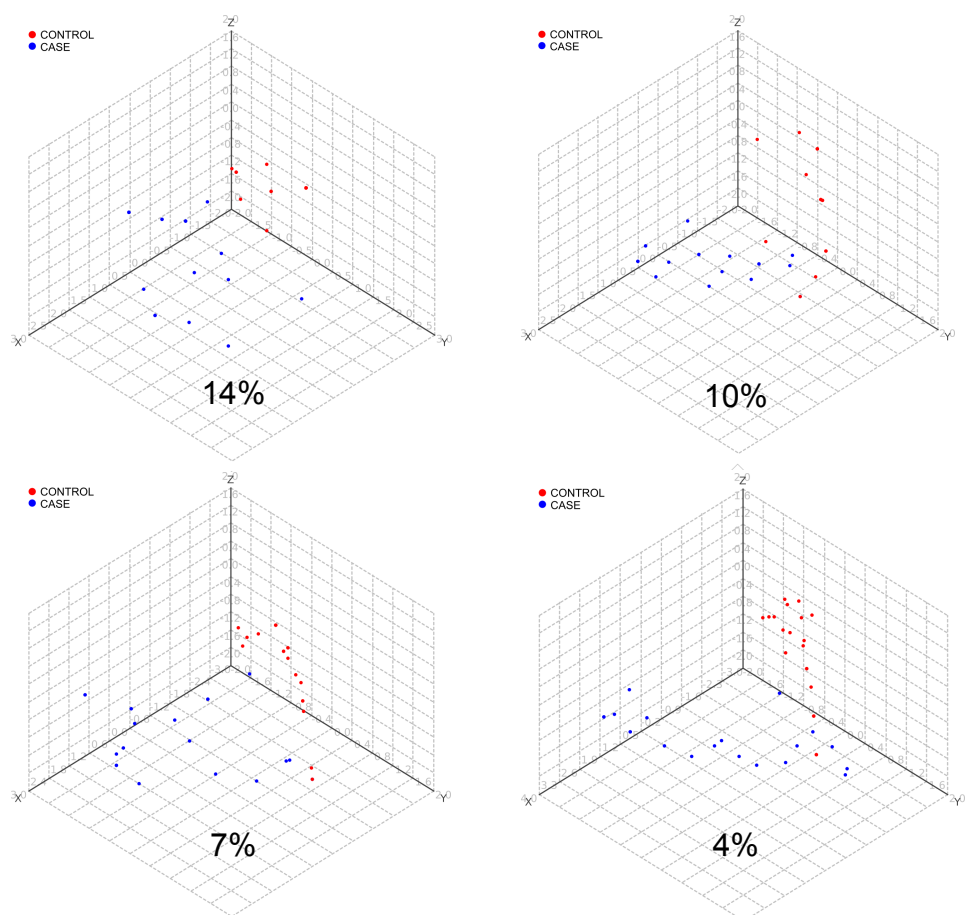


Figura 9. Análisis de componentes principales (PCA) realizado con los péptidos seleccionados de artrosis y sanos para la clasificación de las muestras en estas dos condiciones, en el que 4, 7, 10 y 14 % corresponden a los % de ACN con los que se eluyeron las muestras.

Péptidos seleccionados de una condición					Péptidos comunes a varias condiciones				
m/z	APS	AR	OA	N	m/z	APS	AR	OA	N
1	0%	5%	50%	10%	52	65%	50%	65%	5%
2	5%	15%	40%	0%	53	80%	80%	70%	15%
3	5%	5%	40%	0%	54	65%	85%	70%	20%
4	0%	5%	40%	0%	55	70%	90%	65%	10%
5	10%	5%	45%	0%	56	60%	85%	50%	0%
6	15%	10%	70%	20%	57	60%	50%	55%	5%
7	5%	5%	40%	5%	58	60%	80%	75%	15%
8	5%	10%	60%	20%	59	60%	80%	60%	5%
9	5%	0%	45%	0%	60	45%	45%	50%	0%
10	0%	5%	35%	0%	61	55%	65%	50%	0%
11	35%	25%	90%	40%	62	50%	70%	35%	0%
12	0%	5%	35%	0%	63	75%	60%	45%	0%
13	5%	5%	60%	15%	64	55%	50%	70%	5%
14	0%	0%	0%	30%	65	55%	90%	65%	5%
15	15%	15%	5%	50%	66	95%	90%	75%	20%
16	5%	5%	10%	50%	67	90%	65%	50%	10%
17	0%	0%	15%	55%	68	100%	80%	85%	35%
18	0%	0%	0%	30%	69	80%	95%	80%	45%
19	0%	0%	5%	40%	70	95%	100%	70%	5%
20	0%	0%	5%	50%	71	65%	60%	60%	10%
21	30%	10%	40%	80%	72	75%	85%	80%	30%
22	0%	0%	5%	45%	73	75%	65%	45%	5%
23	30%	40%	30%	90%	74	65%	90%	75%	25%
24	5%	30%	35%	65%	75	80%	85%	75%	20%
25	0%	0%	10%	45%	76	85%	85%	65%	20%
26	10%	10%	15%	55%	77	90%	75%	60%	10%
27	70%	20%	5%	0%	78	80%	90%	80%	20%
28	30%	5%	0%	0%	79	90%	95%	75%	30%
29	40%	10%	0%	5%	80	85%	95%	95%	30%
30	50%	0%	0%	0%	81	70%	80%	70%	5%
31	65%	10%	20%	0%	82	100%	85%	50%	15%
32	50%	20%	5%	0%	83	75%	90%	95%	40%
33	40%	10%	5%	0%	84	75%	60%	55%	15%
34	30%	0%	0%	0%	85	75%	85%	65%	5%
35	85%	40%	0%	0%	86	5%	35%	50%	55%
36	40%	5%	0%	0%	87	10%	45%	60%	60%
37	30%	0%	0%	0%	88	25%	95%	75%	70%
38	60%	30%	20%	0%	89	50%	80%	0%	0%
39	5%	30%	0%	0%	90	45%	45%	10%	0%
40	5%	30%	0%	0%	91	55%	50%	25%	0%
41	5%	30%	0%	0%	92	90%	90%	20%	0%
42	10%	35%	0%	0%	93	50%	70%	35%	0%
43	15%	40%	0%	5%	94	95%	90%	45%	10%
44	0%	45%	10%	0%	95	50%	35%	5%	0%
45	25%	45%	0%	0%	96	75%	55%	20%	0%
46	20%	60%	15%	10%	97	55%	45%	10%	0%
47	20%	40%	0%	0%	98	25%	45%	5%	0%
48	15%	60%	10%	0%	99	95%	90%	25%	20%
49	5%	50%	0%	0%	100	50%	55%	15%	15%
50	10%	35%	0%	0%	101	70%	85%	20%	0%
51	35%	70%	20%	5%	102	100%	75%	15%	0%
					103	5%	30%	80%	80%
					104	0%	5%	80%	90%
					105	0%	5%	80%	90%
					106	25%	25%	70%	65%
					107	0%	45%	45%	10%

Tabla 3. Péptidos seleccionados para la clasificación de las muestras en sanos (N), artrósicos (OA), artritis psoriásica (APS) y artritis reumatoide (AR). Los % corresponden a los porcentajes de presencia de dichos péptidos en las cuatro condiciones.

4. DISCUSIÓN

4. Discusión

4.1. Discusión artículo I

En este trabajo se compararon tres métodos de depleción de sueros que se llevaron a cabo con diferentes compuestos químicos. Lo primero que se realizó fue un estudio de los métodos de cuantificación de proteína disponibles en nuestro laboratorio para escoger el más adecuado para nuestras muestras. Existen diversos métodos de cuantificación basados en la propiedad intrínseca de las proteínas para absorber luz en el ultravioleta (UV), o en la formación de derivados químicos. Se compararon los métodos empleados de rutina en nuestro laboratorio: el de medida de la absorbancia a una longitud de onda de 280 nm con un espectrofotómetro NanoDrop, y los métodos basados en reacciones químicas de Bradford y BCA (ácido bicinconínico). Obtuvimos el mismo resultado en términos de exactitud y precisión con los métodos de nanodrop y BCA, resultando el método de Bradford el más exacto de los tres. Todos estos métodos tienen una serie de ventajas e inconvenientes a considerar: 1) Los métodos de absorción se basan en que los aminoácidos tirosina y triptófano tienen electrones deslocalizados que absorben la radiación ultravioleta, este método es el más rápido, permite la reutilización de las muestras y no necesita añadir ningún reactivo a la muestra, pero interfieren todos los compuestos que puedan absorber a dicha longitud de onda. 2) El método de Bradford se basa en la unión de un colorante (Coomassie) a la proteína, debido a la presencia de aminoácidos básicos con cadenas aromáticas (arginina, lisina e histidina). Es un método muy sensible a la cantidad de proteína presente en la muestra, pero la principal desventaja es su incompatibilidad con los detergentes presentes en algunos tampones de extracción de proteínas. 3) Por último, el método de BCA se basa en la formación de un complejo cuproso con los residuos amino de los péptidos que es detectado por el ácido bicinconínico emitiendo luz azul, es el más sensible y es compatible con la mayoría de los detergentes, sólo pueden interferir sustancias que reduzcan el cobre o que sean quelantes. Teniendo en cuenta esta información y nuestro tipo de muestra, se decidió emplear el nanodrop para cuantificar sueros crudos ya que en este tipo de muestra prácticamente no hay interferencias en la absorción de luz a 280 nm y la sensibilidad era suficiente. En el caso de las muestras deplecionadas, se decidió emplear el método de Bradford por ser el más exacto y debido a que en el

nanodrop se produjeron interferencias con los reactivos químicos presentes en la muestra (como la urea del kit de ProteoMiner).

Antes de comenzar con la comparación de los métodos de depleción, se identificaron las proteínas presentes en las bandas de un gel cargado con suero crudo y se comprobó que en todas las bandas se identificaba la albúmina además de otras proteínas abundantes del suero, dato que reafirma la necesidad de realizar una depleción para poder hacer nuestros estudios proteómicos en suero.

Depleción con ACN

Kay y colaboradores [122] describieron un método simple y eficaz para deplecionar proteínas añadiendo al suero concentraciones de ACN hasta el 60% (v/v). En nuestro laboratorio se comprobó la eficacia de este método realizando un ensayo para escoger la mejor concentración de ACN, determinando (de acuerdo con Kay y colaboradores) que 57% (v/v) es la más adecuada para deplecionar las proteínas abundantes del suero. La precipitación de las proteínas abundantes parece producirse debido a la adición de una sustancia orgánica (ACN) a una disolución acuosa, de manera que se debilitan las interacciones hidrofóbicas y se fortalecen los enlaces peptídicos por puentes de hidrógeno, lo que provoca la desnaturalización de las proteínas. Una vez que se encuentran desnaturalizadas tienden a formar agregados y precipitar [124].

Con nuestra metodología detectamos 21 proteínas, de las cuales 5 pertenecían al grupo de las más abundantes, por lo que se incrementó la concentración de las otras 15 proteínas menos abundantes del suero. Se detectó un elevado porcentaje de apolipoproteínas (29%) en el suero deplecionado, hecho que respalda la hipótesis de Kay y colaboradores en relación a que este método enriquece en proteínas de naturaleza hidrofóbica. Siguiendo este razonamiento, las apolipoproteínas que detectamos probablemente permanecen en solución debido a sus propiedades anfipáticas. Estas proteínas se han relacionado con la artrosis y otras enfermedades reumáticas [125, 126], así como con enfermedades cardiovasculares y con factor de riesgo de infarto de miocardio [127]. Por lo que esta depleción podría ser recomendable para estudiar marcadores de enfermedades reumáticas y cardiovasculares. Por otra parte, se observó en los geles una evidente disminución en la intensidad de la banda correspondiente a la albúmina, lo que demuestra la eficiente depleción de esta proteína con este

método. También se verificó la reproducibilidad del método, ya que el 86% de las proteínas se identificaron en al menos dos de las cuatro réplicas, y en los propios geles se observaron los mismos perfiles proteicos para todas las réplicas.

Depleción con DTT

Warder y colaboradores [121] publicaron un método robusto, rápido y sencillo para deplecionar proteínas abundantes ricas en puentes disulfuro. En el presente trabajo se empleó esta metodología con pequeñas modificaciones, que dio lugar a un perfil proteico del suero en los geles similar al previamente publicado. El DTT se emplea habitualmente para romper los puentes disulfuro tanto intra- como intermoleculares, ayudando a desplegar las proteínas que presentan muchos puentes disulfuro. Una vez desnaturalizadas, tienden a formar agregados que finalmente precipitan. Éste podría ser el mecanismo por el cual la albúmina y serotransferrina entre otras, se deplecionan, ya que son ricas en puentes disulfuro. Al aplicar este método en nuestras muestras se identificaron 38 proteínas, de las cuales 18 pertenecían al grupo de las más abundantes, con lo que en este método se deplecionaron menos que con el de ACN, pero en cambio se obtuvo un mayor número de proteínas identificadas fuera del grupo de las abundantes. En este caso se detectó un porcentaje elevado (37%) de inmunoglobulinas sobre el total de las proteínas identificadas. Estas proteínas están principalmente relacionadas con la respuesta inmune secundaria, haciendo que este método de depleción sea especialmente útil en estudios de enfermedades caracterizadas por modificaciones en el contenido de inmunoglobulinas, como inmunodeficiencias o amiloidosis entre otras [128]. Estas proteínas también se han descrito en artritis reumatoide debido al origen autoinmune de dicha enfermedad [129, 130]. Warder y colaboradores encontraron que las bandas del gel mayores de 95 kDa eran ricas en albúmina, sin embargo, en nuestro trabajo se encontraron otras proteínas (apolipoproteína B-100, complemento C3, α -2-macroglobulina, ceruloplasmina y la cadena pesada del inhibidor inter- α -tripsina H2). Esto puede deberse a las pequeñas modificaciones que se hicieron del protocolo, descritas en la sección de material y métodos del artículo. Warder y colaboradores sugirieron que la depleción con DTT se puede relacionar directamente con el contenido de cisteínas en la proteína. Esta teoría parece coincidir con el hecho de que la serotransferrina y la albúmina tienen una elevada cantidad de cisteínas en relación a su cantidad total de aminoácidos. Sin embargo, esto no ocurre con otras proteínas identificadas en las bandas 9 y 14,

que tienen una proporción similar de cisteínas (glicoproteína- α -2-HS, cadena Ig γ -2 región C y cadena Ig γ -3 región C). Esto puede deberse a la abundancia relativa de dichas proteínas en el suero, ya que la albúmina y la serotransferrina constituyen un 54.3 y 3.3%, respectivamente, de la cantidad de proteína total en el suero, resultando una depleción más eficaz que en el caso de las proteínas identificadas. El 80% de las proteínas identificadas se encontraron en al menos dos de las cuatro réplicas realizadas, de manera que se verificó la reproducibilidad del método. También se pudo comprobar en los geles, la disminución de la banda de albúmina respecto a la del suero crudo, así como un mismo perfil proteico en todas las réplicas.

Ecuilibración de las proteínas con el kit ProteoMiner

El kit ProteoMiner (PM) es una metodología desarrollada por la casa comercial BioRad para reducir y de esta forma poder abarcar analíticamente el rango dinámico de la concentración de proteínas en muestras biológicas complejas, como el suero. La compañía afirma que el empleo de esta tecnología permite la detección de proteínas de abundancia media y baja. El kit se basa en una extensa librería de hexapéptidos unidos a soportes cromatográficos que se mezclan con las muestras. Como la capacidad de unión está limitada, las proteínas más abundantes saturan enseguida sus ligandos y el exceso de estas proteínas se eliminan en el lavado. Por el contrario, las proteínas poco abundantes se concentran en sus respectivos ligandos específicos sin llegar a saturarlos. El resultado es una disminución del rango dinámico de las proteínas en la muestra. En nuestros resultados hemos podido observar esta disminución en los perfiles de los sueros tratados con este kit, dichos perfiles se parecen más al del suero crudo pero con menos cantidad. De hecho, se observa una clara disminución de las bandas de la albúmina y de proteínas presentes a 55 kDa y 26 kDa en comparación con el suero crudo. Al aplicar este método a nuestras muestras, se identificaron 29 proteínas, de las cuales 9 pertenecían al grupo de las abundantes, y en el perfil proteico obtenido se encontraron tanto inmunoglobulinas como apolipoproteínas y otras proteínas no identificadas con los métodos de ACN y DTT. El 80% de las proteínas identificadas se encontró en al menos 2 de las 4 réplicas, también se observaron los mismos perfiles de proteínas en los geles, verificando así la buena reproducibilidad del método.

Comparación de los tres métodos

Con la consecución de este trabajo hemos podido determinar que la depleción con ACN resultó ser la que más proteínas abundantes depleciona, aunque el método de DTT abarca mejor el rango dinámico de la concentración proteica en el suero y permite identificar un mayor número de proteínas comparado con los otros dos métodos.

Cada método proporcionó una información diferente, el ACN enriqueció el suero en apolipoproteínas, el DTT en inmunoglobulinas y el PM en ambos. Sólo 4 proteínas se identificaron con los tres métodos, confirmando que cada uno de ellos aporta datos de una parte distinta del proteoma del suero. Además, el método de ACN favorece la precipitación de proteínas de peso molecular mayor de 72 kDa, mientras que el PM parece que depleciona mejor proteínas de peso menor de 26 kDa.

4.2. Discusión artículo II

El segundo trabajo muestra la aplicación del método de depleción secuencial en estudios proteómicos cuantitativos de búsqueda de biomarcadores de enfermedad. La búsqueda de nuevos biomarcadores es un área en el que está emergiendo la proteómica como una estrategia muy potente para la detección de nuevas proteínas de uso diagnóstico, pronóstico o destinado a la evaluación de terapias. En este campo, se han llevado a cabo un gran número de experimentos proteómicos en los últimos años con el objetivo de caracterizar nuevas moléculas con potencial biomarcador de enfermedades articulares, incluyendo la artrosis (OA) [131]. En el presente trabajo se muestra la combinación de un método de depleción química secuencial combinado con una electroforesis diferencial en gel bidimensional (2D-DIGE) para la búsqueda de biomarcadores de OA. El método de depleción secuencial reduce el rango dinámico de las proteínas en suero, permitiendo así la identificación de un mayor número de proteínas en las muestras.

Depleción secuencial de suero humano para análisis proteómicos

En el anterior artículo [132] (artículo I) se publicó la comparación de dos métodos de depleción química diferentes, usando acetonitrilo (ACN) o ditiotreitól (DTT), con la ecualización de proteínas del kit ProteoMiner [132] para el análisis mediante espectrometría de masas de las proteínas del suero. Nuestros resultados mostraron cómo la depleción con ACN fue eficiente para deplecionar proteínas de peso molecular mayor de 75 kDa, mientras que la depleción con DTT promueve principalmente la precipitación de proteínas ricas en puentes disulfuro (mayoritariamente albúmina). En este nuevo trabajo se combinaron estas dos depleciones químicas de forma secuencial con el objetivo de simplificar aún más el perfil proteómico del suero, eliminando las proteínas más abundantes que normalmente no aportan información relevante. Los sueros se sometieron a una primera depleción con DTT, y el sobrenadante se volvió a deplecionar con el método de ACN para reducir al máximo posible el rango dinámico de la concentración proteica en el suero.

Primero se realizó un estudio para comprobar la reproducibilidad de la depleción secuencial en cinco réplicas técnicas (cinco depleciones independientes del mismo suero) y en tres réplicas biológicas (tres sueros diferentes) obteniendo una alta reproducibilidad del método, ya que prácticamente el 100% de las proteínas se

identificaron en todas las réplicas y el perfil proteico que se obtuvo en los geles fue muy similar.

Se compararon los perfiles proteicos del suero crudo (ND) y deplecionado (D) mediante geles monodimensionales y bidimensionales (2D), pudiéndose observar un claro descenso de la cantidad de albúmina y otras proteínas mayoritarias como la transferrina, la alfa-2-macroglobulina, etc., así como un enriquecimiento en proteínas que estaban enmascaradas por las abundantes en las muestras crudas.

Análisis cuantitativo 2D-DIGE en las muestras deplecionadas secuencialmente

El método de depleción secuencial se aplicó al estudio de proteínas séricas moduladas en OA. Estas proteínas alteradas en muestras de pacientes enfermos podrían ser útiles como biomarcadores de OA para diagnóstico o pronóstico, y también para la evaluación de terapias alternativas.

Con este objetivo se agruparon las muestras de cada condición (10 muestras/grupo) para reducir la variabilidad individual y biológica que se ha descrito en proteínas plasmáticas [133]. Aunque esto implica una pérdida de los datos individuales de cada muestra, la estrategia de agrupar muestras ha demostrado ser útil para búsqueda de biomarcadores en estrategias proteómicas al azar [134]. Con el fin de realizar un análisis de cuantificación relativa, se siguió la estrategia DIGE [108], que hasta ahora es el mejor método cuantitativo basado en gel. Esta aproximación permite una cuantificación directa de los cambios, reduce la variación entre geles y los falsos positivos [135, 136], de manera que los resultados tienen alta confianza y significancia biológica. Otra ventaja del 2D-DIGE frente a técnicas proteómicas cuantitativas no basadas en gel es que puede detectar directamente no sólo cambios de cantidad de proteína, sino que también modificaciones post-traduccionales [137], las cuales son considerablemente abundantes en las muestras de suero.

Las imágenes obtenidas a partir de los geles se analizaron con el programa SameSpots que normaliza la intensidad de las manchas y calcula los ratios y el poder estadístico de los resultados.

Potenciales biomarcadores de artrosis identificados en este trabajo

Este trabajo permitió detectar 46 manchas significativa y reproduciblemente alteradas entre muestras OA y control que correspondieron a 16 formas proteicas diferentes (Tabla 3). Curiosamente, este panel de proteínas difiere sustancialmente del que se obtuvo en un trabajo previo de nuestro grupo basado en inmunodepleción y análisis de cromatografía líquida acoplada a espectrometría de masas en tándem (LC-MS/MS) (Anexo V). En el trabajo mencionado se encontraron varias proteínas bastante abundantes del suero aumentadas en las muestras de pacientes con OA. Sin embargo, la mayoría de las proteínas complemento fueron eliminadas, mientras que se identificaron varias apolipoproteínas. Estos datos demuestran la complementariedad de los diferentes métodos de depleción.

Por otro lado, aunque en el presente estudio mediante DIGE no se identificaron proteínas poco abundantes como moduladas, sí se pudo poner de manifiesto la utilidad de las estrategias basadas en gel para caracterizar modificaciones proteicas asociadas a enfermedades, que no se podrían obtener con el empleo de aproximaciones basadas en espectrometría de masas (en las que no se analizan proteínas intactas, sino péptidos). Esto incluye no sólo la identificación de una isoforma específica de la cadena beta de la haptoglobina (HPT) aumentada en OA, sino también la identificación de diferentes formas de la proteína amiloide A de suero, transtiretina o algunas apolipoproteínas como moduladas en el suero de pacientes diagnosticados con la enfermedad.

No obstante, ambos métodos han permitido la identificación de varias proteínas relacionadas con el transporte de lípidos, respuesta inmune o proteínas de unión. Aunque las moléculas que se han empleado típicamente como biomarcadores de OA son proteínas bien directa o indirectamente implicadas en la degradación de cartílago, o sintetizadas para reparar este tejido (lo que incluye diferentes tipos de fragmentos de colágeno II, así como otros componentes característicos de su matriz extracelular) [138], cada vez hay más evidencias de que la estrategia más prometedora en OA podría ser la combinación de diferentes paneles de biomarcadores. Éstos deberían incluir proteínas altamente específicas de la patología, pero también proteínas que puedan ser alteradas durante la progresión de procesos inespecíficos como la inflamación, la respuesta inmune, la muerte o proliferación celular, etc. El posible valor biomarcador de esas proteínas se puede ver como ejemplo en un trabajo reciente en donde aclaran el papel clave de

algunos componentes del sistema del complemento en la patogénesis de OA [139]. Además, en el presente trabajo se han identificado también proteínas relacionadas con el metabolismo lipídico, que podrían ser relevantes para la definición de un nuevo fenotipo llamado “atrosia metabólica”, que recientemente está adquiriendo una mayor atención en investigación [140].

Modulación del perfil proteico de la haptoglobina en sueros de pacientes artrósicos

Sorprendentemente, en este trabajo se detectaron tres manchas correspondientes al mismo número de acceso proteico, haptoglobina, que presentaron diferentes alteraciones en los sueros OA comparado con los controles. Dos de ellas aumentadas y una disminuida. La haptoglobina es una glicoproteína N-ligada que se rompe en dos cadenas (alfa y beta) [141], se combina con la hemoglobina libre del plasma y contribuye a mantener la homeostasis celular del hierro. Estudios proteómicos previos en fluido sinovial, empleando geles 2D, mostraron cambios destacables en los patrones individuales de haptoglobina de los pacientes estudiados [142, 143]. A diferencia de estos artículos, en el presente trabajo no hemos analizado muestras individuales sino que realizamos agrupaciones (*pool/s*) de 10 muestras, una estrategia que demostró su utilidad en nuestro estudio previo basado en marcajes isotópicos (iTRAQ) (Anexo V), y que en este trabajo ha permitido la detección de los cambios en las cadenas de la HPT con una elevada significación estadística. La modulación de la cadena beta de la HPT se confirmó mediante análisis de inmunodetección de tipo western blot en 30 nuevas muestras (15 de cada condición), también con significación estadística. Aunque se ha descrito la relación entre los polimorfismos de la HPT con enfermedades como el cáncer o la artritis reumatoide debido a sus propiedades antioxidantes [144], anti-inflamatorias [145] e inmuno-modulatorias [146], esta es la primera vez que se describe la modulación de esta proteína en artrosis. Será necesario realizar más estudios con el fin de verificar la alteración de las otras cadenas de HPT, y poder investigar la relevancia biológica de estas modulaciones.

4.3. Discusión artículo III

En este trabajo se estudió el efecto de los ultrasonidos indirectos en la digestión de sueros artrósicos con tripsina inmovilizada en soportes magnéticos para mejorar la identificación de las proteínas mediante espectrometría de masas.

Las condiciones empleadas para distribuir la energía de los ultrasonidos dentro de las muestras para acelerar los procesos de digestión se escogieron en base a nuestra experiencia con los dispositivos de ultrasonidos y su uso en tratamientos de muestras para proteómica [87, 97, 98, 147-149]. Se escogió el 50% de amplitud de las ondas ultrasónicas porque se ha demostrado que este porcentaje ayuda a la digestión proteica cuando se aplica energía de ultrasonidos de alta intensidad de forma directa (DHI-UE) y las proteínas y el enzima se encuentran en disolución [97, 98, 147-149], o cuando el enzima se encuentra inmovilizado en un soporte sólido [87, 150]. Sin embargo, en este trabajo se comprobó que si se sustituyen los ultrasonidos directos (DHI-UE) por la aplicación indirecta de dicha energía (IHI-UE), la eficiencia de la digestión disminuye drásticamente cuando el enzima está inmovilizado. Para este estudio, se utilizó un dispositivo acoplado a la sonda (cup-horn) que permite transmitir los ultrasonidos de forma indirecta, es decir, la energía se transmite a través de un líquido (agua) a los tubos inmersos en él, donde se encuentran las muestras. De esta forma se evita la contaminación que se puede producir al entrar en contacto directo el extremo de la sonda con nuestra muestra. Se realizó una comparación de la digestión acelerada con ultrasonidos indirectos empleando el enzima en solución frente al enzima inmovilizado en un soporte sólido. Para esta comparación decidimos emplear el marcaje isotópico con ^{18}O .

Eficacia de la doble incorporación de ^{18}O

Cuando la digestión de un proteoma se realiza en presencia de agua- ^{18}O (agua pesada), se pueden incorporar uno o dos átomos de oxígeno en el extremo C-terminal de los péptidos. Para asegurar una correcta cuantificación, el tratamiento de la muestra debe de alcanzar una doble incorporación de ^{18}O superior al 90% [147]. Esto se debe a que si el doble marcaje con ^{18}O no ha sido completo podría ocurrir un solapamiento de los picos isotópicos cuando la muestra marcada y no marcada se mezclan (ver figura 1 del material suplementario del artículo en el anexo III). La comparación entre tripsina inmovilizada y tripsina en solución se

realizó en tres muestras de suero independientes. Cada muestra la dividimos en dos partes iguales que se sometieron a los dos tipos de digestión estudiados, la parte digerida con la tripsina inmovilizada se marcó posteriormente con ^{18}O y cada muestra se analizó por cuádruplicado en el espectrómetro de masas. Para realizar la comparación, primero se analizaron las dos condiciones por separado para escoger sólo los péptidos comunes a la muestra marcada y sin marcar, que además cumpliesen una doble incorporación del ^{18}O mayor del 90% para evitar así el problema anteriormente mencionado.

Influencia del tipo de tripsina en la eficacia de la digestión

La eficacia de la digestión en función de la tripsina empleada se evaluó por la relación entre la intensidad del péptido marcado y la intensidad del péptido sin marcar, para cada péptido presente en las dos condiciones (ligero y pesado) y con una eficacia del doble marcaje superior al 90%. En las tres muestras se obtuvo un ratio mayor del 1.2 en más del 60% de los péptidos analizados, por lo que la digestión con tripsina en solución resultó más eficaz que la tripsina inmovilizada. Además, se encontró un gran número de péptidos que aparecían cuando se analizaron las dos condiciones por separado, pero que cuando se realizó la mezcla 1:1 de ambas, sólo se encontraron en la fracción marcada. En estos péptidos la digestión con tripsina en solución resultó tan eficaz en comparación con la digestión con tripsina inmovilizada que la cuantificación no fue posible, al no poder detectar la versión no marcada.

Influencia del tipo de ultrasonidos aplicado

Teniendo en cuenta los resultados obtenidos, resultó evidente que cuando se emplean los ultrasonidos indirectos para acelerar la digestión enzimática, no se debe utilizar la tripsina inmovilizada en soportes sólidos. Nuestra hipótesis para explicar este hecho es que cuando la tripsina está inmovilizada en un soporte sólido tiende a formar agregados cuando se aplican los ultrasonidos indirectos. En cambio, al aplicar ultrasonidos directos (sonda) no se observa ningún tipo de agregado. Este hecho se puede comprobar en los vídeos 1 y 2 del material suplementario del artículo (en la versión electrónica). Para corroborar esta teoría, se realizó otro experimento en donde se compararon los dos tipos de sonicación mediante un nuevo marcaje con ^{18}O y se calculó el ratio de la misma forma que en la comparación anterior. La eficacia de la digestión fue mayor con la aplicación de

ultrasonidos directos en el 70% de los casos, lo que confirmó la hipótesis de que la disminución de la eficacia de la digestión con la tripsina inmovilizada fue debida a los agregados que se forman al aplicar ultrasonidos indirectos; la tripsina no se homogeniza bien en la muestra, de forma que no consigue digerir todo el contenido de proteína.

4.4. Discusión capítulo IV

La digestión enzimática de un proteoma complejo genera un gran número de péptidos. Estos péptidos se pueden separar de forma selectiva en función de las interacciones que establecen con una fase estacionaria como la C18 en las NuTips, que está formada por cadenas de hidrocarburos de 18 carbonos. Este tipo de fases estacionarias son de carácter apolar e interaccionan preferentemente con péptidos hidrofóbicos, y se suelen conocer con el nombre de fase reversa. En este trabajo, este tipo de separación se aplicó a las muestras de suero deplecionadas y digeridas, optimizando diferentes variables que afectaban al tratamiento de los sueros.

Perfiles proteicos con la depleción con DTT

Cuando se trabaja con dispositivos de extracción en fase sólida, como las NuTips empleadas en este trabajo, se debe tener presente la capacidad de carga de los mismos para evitar saturarla con los péptidos que se unen a la resina. En el caso de las NuTips que se emplearon, tienen una capacidad de carga máxima de 25 µg. Este problema dificulta el estudio de los péptidos por espectrometría de masas cuando se emplea una fuente de ionización de desorción por laser asistida por matriz (MALDI), porque la presencia de algunos péptidos a altas concentraciones podría producir la supresión de señales de los menos abundantes. Esto significa que el espectro obtenido en esta situación tendría unos pocos picos dominando todo el espectro, algo que no es conveniente si el objetivo es encontrar diferencias comparando los picos m/z obtenidos de espectros de diferentes muestras. Para estudiar la influencia de la cantidad de proteína tratada, se realizó un primer experimento con sueros deplecionados con DTT y digeridos, en el que se probaron dos cantidades para cargar en las NuTips, una que sobrepasaba muy poco el límite de carga (30 µg) y otra por debajo de dicho límite (18 µg). Los resultados mostraron un menor número de proteínas identificadas cuando se superaba la capacidad de carga, por lo que se concluyó que la cantidad óptima de proteína se encontraba en torno a 18 µg. Esto demuestra que se debe trabajar por debajo de la capacidad máxima de carga si se quiere evitar perder información. Se debe tener en cuenta que las proteínas se identificaron mediante fragmentación peptídica (MS/MS) de una gota de muestra en placa MALDI, esto implica que el número de

proteínas identificadas es muy pequeño en comparación con un método de separación cromatográfica (HPLC o LC-MALDI).

En este experimento también se observó que el número de proteínas identificadas aumentó hasta el 35% (v/v) de ACN, por lo que a mayores % no eluyeron nuevos péptidos y no se consiguió más información. De acuerdo con estos resultados se probó a cambiar el gradiente de ACN a un nuevo rango de concentraciones más bajas, empleando la cantidad de proteína optimizada anteriormente, pero el número total de proteínas identificadas fue similar al anterior y no se consiguió aumentar la información obtenida.

Se obtuvo una reproducibilidad del método de separación secuencial con NuTips mayor del 80% en cinco réplicas técnicas, por lo que este método de separación resultó ser robusto y reproducible. Cabe destacar que más de un 50% de las proteínas identificadas en estos primeros experimentos pertenecían al grupo de las 20 proteínas más abundantes del suero (Top 20), por lo que se modificó la depleción para evitar la elevada presencia de dichas proteínas.

Perfiles proteicos con la depleción química secuencial

En estudios previos realizados por nuestro grupo [151] (artículo II), se describió un nuevo método de depleción basado en dos precipitaciones químicas consecutivas, que proporciona un suero con un contenido menor en proteínas abundantes, incrementando el número de picos en el espectro MALDI. De acuerdo con estos resultados previos, se realizó un nuevo experimento en el que se escogió este método de depleción química secuencial para sustituir al de DTT y disminuir así el % tan elevado de proteínas abundantes obtenido en los experimentos anteriores del presente trabajo. El número total de proteínas identificadas resultó ser similar, pero en este caso sólo el 22% pertenecían al grupo de las Top 20, consiguiendo de este modo reducir la presencia de proteínas abundantes. También se observó que sólo 5 proteínas se identificaron con ambos métodos. La reproducibilidad del método fue similar a la de la estrategia anterior con la depleción con DTT. Para la optimización del método final de búsqueda de biomarcadores se escogió la depleción química secuencial con el fin de evitar en la medida de lo posible la identificación de proteínas abundantes, y también se simplificó el gradiente de ACN quedándonos principalmente con las fracciones que mayor información aportaron (0%, 4%, 7%, 14% y 35%).

Perfiles proteicos diferenciales entre sueros de pacientes con artrosis y controles sanos

El método final desarrollado a partir de los resultados de los anteriores experimentos se aplicó a la búsqueda de biomarcadores de artrosis. Para ello se emplearon muestras de sueros de 10 pacientes con artrosis (OA) y 10 muestras control. Estas muestras se agruparon en 5 sueros por grupo y por condición, de forma que obtuvimos dos muestras finales de OA y dos de controles. Cada grupo de muestras se sometió a la depleción secuencial, seguida de la digestión con tripsina acelerada por ultrasonidos y a la separación por elución secuencial descrita en el presente trabajo. Los péptidos identificados en cada fracción mostraron que el patrón de proteínas identificadas en cada elución resultó diferente entre los sueros procedentes de pacientes enfermos y los procedentes de donantes sanos.

Clasificación de muestras de suero a partir de su perfil peptídico

Como el resultado anterior mostró que cada fracción eluída proporcionó un patrón característico de proteínas dentro del mismo grupo de donantes, se planteó dar un paso más y comparar los perfiles de los péptidos (no proteínas) obtenidos en cada fracción. Las extracciones en fase sólida con resina de C18 se suelen emplear para preconcentrar y desalar la muestra antes de analizarla por espectrometría de masas (MS) con una única elución [152-154], pero en este trabajo se describe por primera vez la depleción química secuencial acoplada a una separación de péptidos por elución secuencial usando puntas con C18 inmovilizado para reducir el rango dinámico de la concentración proteica y fraccionar las muestras muy complejas, como el suero humano digerido. Para comprobar la utilidad de este método para la búsqueda de biomarcadores de enfermedades reumáticas, se llevó a cabo otro experimento con un total de 80 muestras individuales divididas en 20 sueros de pacientes con artrosis (OA), 20 sueros de pacientes con artritis psoriásica (APS), 20 sueros de pacientes con artritis reumatoide (AR) y 20 sueros de donantes control (N). Se trataron de forma individual con la metodología optimizada: se deplecionaron secuencialmente, se digirieron y se fraccionaron con la elución secuencial, para finalmente ser analizados por quintuplicado en el espectrómetro de masas. Los péptidos que se obtuvieron en las muestras de artrosis y controles se emplearon para realizar un análisis de componentes principales (PCA) con las 4 eluciones realizadas (4, 7, 10 y 14% de ACN), obteniendo la clasificación en al menos dos planos. Esto demuestra la utilidad de

los péptidos para diferenciar muestras de suero de donantes artrósicos de los sueros de donantes controles. Con los péptidos de los espectros representativos de las cuatro condiciones estudiadas se elaboró una lista de 107 péptidos que demostraron tener capacidad para diferenciar entre las cuatro condiciones (OA, N, APS y AR). Dentro de los 107 se encontraron 34 péptidos comunes a las enfermedades reumáticas que tenían una frecuencia de presencia significativamente mayor en los sueros de dichas enfermedades con respecto a los sueros de los controles, lo que indica un perfil peptídico capaz de diferenciar entre estos dos tipos de muestras. Igualmente, se detectaron 13 péptidos presentes con mayor frecuencia en sueros de donantes artrósicos con respecto a los sueros de pacientes con las otras dos enfermedades reumáticas (APS y AR) y a los sueros controles (N). Estos péptidos se consideraron característicos de la patología artrósica, ya que no se detectaron en otras patologías reumáticas ni en los controles sanos. Finalmente, también se detectaron otros grupos de péptidos que se consideraron característicos de las otras patologías, así como de las muestras control.

Estos resultados demuestran la utilidad de la metodología desarrollada como estrategia de flujo de trabajo en estudios proteómicos para la búsqueda de perfiles peptídicos que permitan clasificar las muestras.

4.5. Consideraciones finales

La búsqueda de biomarcadores moleculares de enfermedad está siendo una prioridad en la investigación de estos últimos años, debido a que hay un gran número de patologías que no se pueden diagnosticar a tiempo de evitar daños irreparables en el paciente (como ocurre en la artrosis), y también a las dificultades que existen para poder evaluar nuevos tratamientos. Con este objetivo de descubrimiento de nuevos marcadores, se han desarrollado numerosas estrategias proteómicas que caracterizan proteínas del suero como potenciales biomarcadores de enfermedades [52, 126, 155-160]. Muchas de estas estrategias constan de una depleción con columnas cromatográficas de inmutafinidad, una digestión con tripsina, una separación de los péptidos mediante uno o dos tipos de cromatografía líquida, y un análisis de los péptidos mediante espectrometría de masas con el posterior análisis bioinformático de los datos. Esta metodología produce una gran cantidad de información que en muchos casos es complicada de interpretar, haciendo muy difícil escoger unas proteínas o péptidos con un claro potencial biomarcador. Además, tiene como inconveniente añadido el alto coste de las mismas y el tiempo que demoran. Las técnicas cromatográficas necesitan una infraestructura de coste elevado, tanto en los aparatos de alta eficacia (HPLC) o de nanocromatografía líquida (nanoLC), como en las columnas cromatográficas empleadas. Este tipo de metodología suele alargar bastante el tratamiento de las muestras, además de necesitar un mantenimiento específico que también conlleva un consumo de tiempo extra.

El objetivo de este trabajo fue encontrar una metodología proteómica alternativa que permita evitar todos estos inconvenientes, de manera que las muestras de suero se puedan tratar de forma rápida, sencilla y económica, proporcionando la suficiente información que permita encontrar posibles biomarcadores de enfermedades reumáticas.

En el primer trabajo, hemos comparado dos métodos de depleción sencillos y rápidos, basados en precipitaciones químicas, con un kit comercial diseñado para el tratamiento específico de sueros. Con él demostramos que los dos métodos de depleción química son útiles para reducir el rango dinámico de la concentración proteica del suero, y que proporcionan diferente información del suero: el método

con ACN enriquece en apolipoproteínas mientras que el método con DTT enriquece en inmunoglobulinas. Estos resultados son útiles a la hora de escoger un método de depleción, ya que hay patologías reumáticas para las que se ha descrito relación con determinados tipos de proteínas. También se debe considerar la posibilidad de aplicar los dos tipos de depleción por separado, si se quiere obtener toda la información posible del suero mediante este tipo de depleciones.

Basándonos en los resultados anteriores, se realizó un segundo trabajo en el que se desarrolló un nuevo método de depleción. Este método consiste en una depleción química secuencial, primero aplicando la precipitación con DTT y después con ACN. De esta manera, demostramos una mayor disminución de las proteínas abundantes del suero que con las depleciones por separado. Se comprobó la utilidad del método en una búsqueda de biomarcadores de artrosis realizada mediante la técnica DIGE, que permitió definir un panel de 16 formas proteicas con potencial biomarcador. También se validó en este trabajo por primera vez el aumento de la cadena beta de la haptoglobina en pacientes artrósicos. Esta proteína había sido relacionada con numerosas patologías diversas, pero nunca antes con la artrosis [161-166].

Hemos descrito anteriormente la aplicación de los ultrasonidos de alta intensidad para acelerar el proceso de digestión de proteínas, comprobando que si se aplican de forma indirecta sobre las muestras se evitan las contaminaciones producidas por el contacto entre la sonda y las mismas, y además se acortan los tiempos ya que se puede acelerar la digestión de varias muestras a la vez. Teniendo en cuenta estos resultados, realizamos nuestro tercer trabajo con el fin de aplicar los ultrasonidos indirectos en la digestión de nuestras muestras, pero empleando el enzima inmovilizado en un soporte sólido para tratar de simplificar los espectros de masas al evitar las interferencias de los picos de tripsina. Nuestros resultados demostraron que no se obtiene una buena digestión de las proteínas con la tripsina inmovilizada si se aplican ultrasonidos indirectos, ya que éstos perjudican la buena homogenización del enzima. Por ello, en el desarrollo de nuestra nueva metodología proteómica se escogieron los ultrasonidos indirectos con el enzima en solución para el paso de digestión proteica.

Por último, desarrollamos un método de separación peptídica sencillo basado en eluciones secuenciales de los péptidos retenidos en una extracción en fase sólida de C18 (en puntas tipo NuTips) con diferentes concentraciones de acetonitrilo. En este último trabajo se optimizó la estrategia completa para la búsqueda de biomarcadores mediante espectrometría de masas, basada en la depleción química secuencial, seguida de la digestión acelerada por ultrasonidos indirectos y por último la separación por elución secuencial en puntas con C18. Las fracciones obtenidas fueron analizadas mediante espectrometría de masas, realizando a continuación un análisis informático con los datos que se extrajeron a partir de los espectros de masas adquiridos. Con este trabajo hemos podido demostrar la utilidad de este flujo de trabajo en la búsqueda de biomarcadores, al encontrar perfiles peptídicos que diferencian sueros de pacientes con enfermedades reumáticas de sueros de pacientes sin historia de patología articular.

5. CONCLUSIONES

5. Conclusiones

1. El estudio comparativo de los métodos de depleción por precipitación química con DTT o ACN muestra una clara disminución de las proteínas más abundantes del suero, resultando un método útil para estudiar muestras de suero desde el punto de vista proteómico.
2. La precipitaciones con DTT y ACN proporcionan un perfil proteico diferente. La depleción con DTT enriquece el suero en inmunoglobulinas, mientras que la de ACN lo enriquece en apolipoproteínas y, en general, en proteínas de peso molecular menor de 75 kDa.
3. El nuevo método de depleción desarrollado, la depleción química secuencial, permite una mayor disminución de las proteínas abundantes del suero en comparación con las dos precipitaciones por separado.
4. La aplicación de la depleción secuencial al estudio comparativo entre sueros de pacientes artrósicos y sueros de donantes sin síntomas mediante la técnica DIGE, proporcionó un panel de 16 formas proteicas (A1AT, FETUA, ANGT, APOA1, APO2, APOA4, APOC3, HPT cadena β , HPT cadena α -1, HPT cadena α -2, HBB, IGKC, LAC2, A2GL, SAA y TTHY) moduladas en la patología artrósica.
5. La cadena beta de la haptoglobina se encuentra en mayor cantidad en sueros de pacientes con artrosis que en los controles sanos. En base a estos resultados proponemos esta forma proteica como potencial biomarcador de la artrosis.
6. La digestión con tripsina inmovilizada en un soporte sólido y acelerada por la aplicación de ultrasonidos indirectos de alta intensidad no mejora identificación de las proteínas, en comparación con la tripsina en solución bajo las mismas condiciones.
7. La estrategia final desarrollada, basada en la depleción secuencial, la digestión acelerada por ultrasonidos indirectos, la separación de los péptidos mediante su elución secuencial del soporte de C18 y el análisis mediante espectrometría de masas tipo MALDI-TOF, es un método sencillo y barato para descubrir

biomarcadores potencialmente útiles para el diagnóstico de la artrosis, la artritis psoriásica y la artritis reumatoide.

8. La aplicación de la estrategia descrita en el punto 7 nos ha permitido validar en suero de pacientes un perfil proteómico formado por 107 péptidos con potencial biomarcador de tres enfermedades reumáticas: OA, AR y APS.

6. BIBLIOGRAFÍA

6. Bibliografía

- [1] Firestein, G. S., Budd, R. C., Jr., E. D. H., McInnes, I. B., *et al.*, *Kelley's Textbook of Rheumatology*, Saunders Elsevier 2009.
- [2] Benito-Ruiz, P., García, F. J. B., Molina, J. T., Parera, I. M., *et al.*, 2010, p. 65.
- [3] FJ., B., JL., F.-S., *Reumatología: Enfermedades del Aparato Locomotor*, Editorial Aran 1998, pp. 1609-1642.
- [4] Woolf, A. D., Pfleger, B., Burden of major musculoskeletal conditions. *Bull World Health Organ* 2003, **81**, 646-656.
- [5] Carmona, L., Ballina, J., Gabriel, R., Laffon, A., Group, E. S., The burden of musculoskeletal diseases in the general population of Spain: results from a national survey. *Ann Rheum Dis* 2001, **60**, 1040-1045.
- [6] Consumo, I. N. d. E. y. M. d. S. y., in: Consumo, I. N. d. E. y. M. d. S. y. (Ed.) 2006.
- [7] Benito-Ruiz, P., García, F. J. B., Molina, J. T., Parera, I. M., *et al.*, *Manual de la SER: ARTROSIS. Fisiopatología, diagnóstico y tratamiento*, 2010.
- [8] Benito-Ruiz, P., García, F. J. B., Molina, J. T., Parera, I. M., *et al.*, 2010, pp. 31-35.
- [9] Dunlop, D. D., Manheim, L. M., Yelin, E. H., Song, J., Chang, R. W., The costs of arthritis. *Arthritis Rheum* 2003, **49**, 101-113.
- [10] Loza, E., Lopez-Gomez, J. M., Abasolo, L., Maese, J., *et al.*, Economic burden of knee and hip osteoarthritis in Spain. *Arthritis Rheum* 2009, **61**, 158-165.
- [11] Felson, D. T., Lawrence, R. C., Dieppe, P. A., Hirsch, R., *et al.*, Osteoarthritis: new insights. Part 1: the disease and its risk factors. *Ann Intern Med* 2000, **133**, 635-646.
- [12] KELLGREN, J. H., Osteoarthritis in patients and populations. *Br Med J* 1961, **2**, 1-6.
- [13] Zhang, Y., McAlindon, T. E., Hannan, M. T., Chaisson, C. E., *et al.*, Estrogen replacement therapy and worsening of radiographic knee osteoarthritis: the Framingham Study. *Arthritis Rheum* 1998, **41**, 1867-1873.
- [14] Wluka, A. E., Davis, S. R., Bailey, M., Stuckey, S. L., Cicuttini, F. M., Users of oestrogen replacement therapy have more knee cartilage than non-users. *Ann Rheum Dis* 2001, **60**, 332-336.
- [15] Loughlin, J., The genetic epidemiology of human primary osteoarthritis: current status. *Expert Rev Mol Med* 2005, **7**, 1-12.
- [16] Valdes, A. M., McWilliams, D., Arden, N. K., Doherty, S. A., *et al.*, Involvement of different risk factors in clinically severe large joint osteoarthritis according to the presence of hand interphalangeal nodes. *Arthritis Rheum* 2010, **62**, 2688-2695.
- [17] Valdes, A. M., Evangelou, E., Kerkhof, H. J., Tamm, A., *et al.*, The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Ann Rheum Dis* 2011, **70**, 873-875.
- [18] Evangelou, E., Valdes, A. M., Kerkhof, H. J., Stykarsdottir, U., *et al.*, Meta-analysis of genome-wide association studies confirms a susceptibility locus for knee osteoarthritis on chromosome 7q22. *Ann Rheum Dis* 2011, **70**, 349-355.
- [19] Day-Williams, A. G., Southam, L., Panoutsopoulou, K., Rayner, N. W., *et al.*, A variant in MCF2L is associated with osteoarthritis. *Am J Hum Genet* 2011, **89**, 446-450.
- [20] Consortium, a., Collaborators, a., Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet* 2012, **380**, 815-823.

- [21] Rego-Pérez, I., Fernández-Moreno, M., Fernández-López, C., Arenas, J., Blanco, F. J., Mitochondrial DNA haplogroups: role in the prevalence and severity of knee osteoarthritis. *Arthritis Rheum* 2008, **58**, 2387-2396.
- [22] Blanco, F. J., Rego, I., Ruiz-Romero, C., The role of mitochondria in osteoarthritis. *Nat Rev Rheumatol* 2011, **7**, 161-169.
- [23] Fernández-Moreno, M., Tamayo, M., Soto-Hermida, A., Mosquera, A., *et al.*, mtDNA haplogroup J modulates telomere length and nitric oxide production. *BMC Musculoskelet Disord* 2011, **12**, 283.
- [24] Fernández-Tajes, J., Soto-Hermida, A., Vázquez-Mosquera, M. E., Cortés-Pereira, E., *et al.*, Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. *Ann Rheum Dis* 2013.
- [25] Rego-Perez, I., Fernandez-Moreno, M., Soto-Hermida, A., Fenandez-Lopez, C., *et al.*, Mitochondrial genetics and osteoarthritis. *Front Biosci (Schol Ed)* 2013, **5**, 360-368.
- [26] Manninen, P., Riihimäki, H., Heliövaara, M., Mäkelä, P., Overweight, gender and knee osteoarthritis. *Int J Obes Relat Metab Disord* 1996, **20**, 595-597.
- [27] Felson, D. T., Zhang, Y., Hannan, M. T., Naimark, A., *et al.*, Risk factors for incident radiographic knee osteoarthritis in the elderly: the Framingham Study. *Arthritis Rheum* 1997, **40**, 728-733.
- [28] Dougados, M., Gueguen, A., Nguyen, M., Thiesce, A., *et al.*, Longitudinal radiologic evaluation of osteoarthritis of the knee. *J Rheumatol* 1992, **19**, 378-384.
- [29] Schouten, J. S., van den Ouweland, F. A., Valkenburg, H. A., A 12 year follow up study in the general population on prognostic factors of cartilage loss in osteoarthritis of the knee. *Ann Rheum Dis* 1992, **51**, 932-937.
- [30] Felson, D. T., Naimark, A., Anderson, J., Kazis, L., *et al.*, The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis Rheum* 1987, **30**, 914-918.
- [31] Gelber, A. C., Hochberg, M. C., Mead, L. A., Wang, N. Y., *et al.*, Joint injury in young adults and risk for subsequent knee and hip osteoarthritis. *Ann Intern Med* 2000, **133**, 321-328.
- [32] Buckwalter, J. A., Lane, N. E., Athletics and osteoarthritis. *Am J Sports Med* 1997, **25**, 873-881.
- [33] Felson, D. T., Hannan, M. T., Naimark, A., Berkeley, J., *et al.*, Occupational physical demands, knee bending, and knee osteoarthritis: results from the Framingham Study. *J Rheumatol* 1991, **18**, 1587-1592.
- [34] P., B.-R., 2002, pp. 33-42.
- [35] Jenniskens, Y. M., Koevoet, W., de Bart, A. C., Weinans, H., *et al.*, Biochemical and functional modulation of the cartilage collagen network by IGF1, TGFbeta2 and FGF2. *Osteoarthritis Cartilage* 2006, **14**, 1136-1146.
- [36] van den Berg, W. B., van Riel, P. L., Uncoupling of inflammation and destruction in rheumatoid arthritis: myth or reality? *Arthritis Rheum* 2005, **52**, 995-999.
- [37] Fernandes, J. C., Martel-Pelletier, J., Pelletier, J. P., The role of cytokines in osteoarthritis pathophysiology. *Biorheology* 2002, **39**, 237-246.
- [38] Loeser, R. F., Molecular mechanisms of cartilage destruction: mechanics, inflammatory mediators, and aging collide. *Arthritis Rheum* 2006, **54**, 1357-1360.
- [39] Monfort, J., Tardif, G., Reboul, P., Mineau, F., *et al.*, Degradation of small leucine-rich repeat proteoglycans by matrix metalloprotease-13: identification of a new biglycan cleavage site. *Arthritis Res Ther* 2006, **8**, R26.
- [40] Stöve, J., Huch, K., Günther, K. P., Scharf, H. P., Interleukin-1beta induces different gene expression of stromelysin, aggrecan and tumor-necrosis-factor-

- stimulated gene 6 in human osteoarthritic chondrocytes in vitro. *Pathobiology* 2000, 68, 144-149.
- [41] Cillero-Pastor, B., Ruiz-Romero, C., Caramés, B., López-Armada, M. J., Blanco, F. J., Proteomic analysis by two-dimensional electrophoresis to identify the normal human chondrocyte proteome stimulated by tumor necrosis factor alpha and interleukin-1beta. *Arthritis Rheum* 2010, 62, 802-814.
- [42] Lee, A. S., Ellman, M. B., Yan, D., Kroin, J. S., *et al.*, A current review of molecular mechanisms regarding osteoarthritis and pain. *Gene* 2013, 527, 440-447.
- [43] P., B.-R., 2002, pp. 91-106.
- [44] Altman, R., Alarcón, G., Appelrouth, D., Bloch, D., *et al.*, The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. *Arthritis Rheum* 1990, 33, 1601-1610.
- [45] Altman, R., Alarcón, G., Appelrouth, D., Bloch, D., *et al.*, The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991, 34, 505-514.
- [46] Altman, R., Asch, E., Bloch, D., Bole, G., *et al.*, Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986, 29, 1039-1049.
- [47] Benito-Ruiz, P., García, F. J. B., Molina, J. T., Parera, I. M., *et al.*, 2010, p. 5.
- [48] Attur, M., Krasnokutsky-Samuels, S., Samuels, J., Abramson, S. B., Prognostic biomarkers in osteoarthritis. *Curr Opin Rheumatol* 2013, 25, 136-144.
- [49] Zhai, G., Eshghi, E. A., Biomarkers for osteoarthritis: investigation, identification, and prognosis. *Current Biomarker Findings* 2012, 2, 19-28.
- [50] Rousseau, J. C., Delmas, P. D., Biological markers in osteoarthritis. *Nat Clin Pract Rheumatol* 2007, 3, 346-356.
- [51] Kraus, V. B., Burnett, B., Coindreau, J., Cottrell, S., *et al.*, Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis. *Osteoarthritis Cartilage* 2011, 19, 515-542.
- [52] Mobasheri, A., Osteoarthritis year 2012 in review: biomarkers. *Osteoarthritis Cartilage* 2012, 20, 1451-1464.
- [53] Ruiz-Romero, C., Blanco, F. J., *Use of biomarkers for the purposes of diagnosis and drug discovery programs: where do we stand?*, Future Medicine Ltd., London 2013, pp. 54-68.
- [54] Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., *et al.*, Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* 1995, 16, 1090-1094.
- [55] Covos, J. V., in: Centro de Biología Molecular Severo Ochoa (CBM), C.-U. (Ed.).
- [56] Ruiz-Romero, C., *Manual de la SER: Artrosis. Fisiopatología, diagnóstico y tratamiento* 2010, pp. 146-162.
- [57] Group., B. D. W., Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001, 69, 89-95.
- [58] Food and Drug Administration, F., in: Services, U. S. D. o. H. a. H. (Ed.) 2008.
- [59] Williams, A., Smith, J. R., Allaway, D., Harris, P., *et al.*, Applications of proteomics in cartilage biology and osteoarthritis research. *Front Biosci (Landmark Ed)* 2011, 16, 2622-2644.
- [60] Etzioni, R., Urban, N., Ramsey, S., McIntosh, M., *et al.*, The case for early detection. *Nat Rev Cancer* 2003, 3, 243-252.
- [61] Rifai, N., Gillette, M. A., Carr, S. A., Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* 2006, 24, 971-983.

- [62] Anderson, N. L., Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 2002, 1, 845-867.
- [63] Schiess, R., Wollscheid, B., Aebersold, R., Targeted proteomic strategy for clinical biomarker discovery. *Mol Oncol* 2009, 3, 33-44.
- [64] Roche, S., Tiers, L., Provansal, M., Piva, M. T., Lehmann, S., Interest of major serum protein removal for Surface-Enhanced Laser Desorption/Ionization - Time Of Flight (SELDI-TOF) proteomic blood profiling. *Proteome Sci* 2006, 4, 20.
- [65] Fountoulakis, M., Juranville, J. F., Jiang, L., Avila, D., *et al.*, Depletion of the high-abundance plasma proteins. *Amino Acids* 2004, 27, 249-259.
- [66] Kovács, A., Guttman, A., Medicinal chemistry meets proteomics: fractionation of the human plasma proteome. *Curr Med Chem* 2013, 20, 483-490.
- [67] Echan, L. A., Tang, H. Y., Ali-Khan, N., Lee, K., Speicher, D. W., Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* 2005, 5, 3292-3303.
- [68] Shen, Y., Kim, J., Strittmatter, E. F., Jacobs, J. M., *et al.*, Characterization of the human blood plasma proteome. *Proteomics* 2005, 5, 4034-4045.
- [69] Anderson, N. L., The roles of multiple proteomic platforms in a pipeline for new diagnostics. *Mol Cell Proteomics* 2005, 4, 1441-1444.
- [70] Cox, J., Mann, M., Is proteomics the new genomics? *Cell* 2007, 130, 395-398.
- [71] T.D., V., J.R., Y., *Proteomics for biological discovery*, New York: John Wiley & Sons 2006.
- [72] Liotta, L. A., Petricoin, E. F., Mass spectrometry-based protein biomarker discovery: solving the remaining challenges to reach the promise of clinical benefit. *Clin Chem* 2010, 56, 1641-1642.
- [73] Ruiz-Romero, C., Blanco, F. J., The role of proteomics in osteoarthritis pathogenesis research. *Curr Drug Targets* 2009, 10, 543-556.
- [74] Bencharit, S., Border, M. B., Edelmann, A., Byrd, W. C., Update in research and methods in proteomics and bioinformatics. *Expert Rev Proteomics* 2013, 10, 413-415.
- [75] Altelaar, A. F., Munoz, J., Heck, A. J., Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat Rev Genet* 2013, 14, 35-48.
- [76] Van Riper, S. K., de Jong, E. P., Carlis, J. V., Griffin, T. J., Mass spectrometry-based proteomics: basic principles and emerging technologies and directions. *Adv Exp Med Biol* 2013, 990, 1-35.
- [77] Aebersold, R., Mann, M., Mass spectrometry-based proteomics. *Nature* 2003, 422, 198-207.
- [78] Jaros, J. A., Guest, P. C., Bahn, S., Martins-de-Souza, D., Affinity depletion of plasma and serum for mass spectrometry-based proteome analysis. *Methods Mol Biol* 2013, 1002, 1-11.
- [79] Urbas, L., Brne, P., Gabor, B., Barut, M., *et al.*, Depletion of high-abundance proteins from human plasma using a combination of an affinity and pseudo-affinity column. *J Chromatogr A* 2009, 1216, 2689-2694.
- [80] Mahn, A., Reyes, A., Zamorano, M., Cifuentes, W., Ismail, M., Depletion of highly abundant proteins in blood plasma by hydrophobic interaction chromatography for proteomic analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* 2010, 878, 1038-1044.
- [81] Holewinski, R. J., Jin, Z., Powell, M. J., Maust, M. D., Van Eyk, J. E., A fast and reproducible method for albumin isolation and depletion from serum and cerebrospinal fluid. *Proteomics* 2013, 13, 743-750.
- [82] Zolotarjova, N., Mrozinski, P., Chen, H., Martosella, J., Combination of affinity depletion of abundant proteins and reversed-phase fractionation in proteomic analysis of human plasma/serum. *J Chromatogr A* 2008, 1189, 332-338.

- [83] Beseme, O., Fertin, M., Drobecq, H., Amouyel, P., Pinet, F., Combinatorial peptide ligand library plasma treatment: Advantages for accessing low-abundance proteins. *Electrophoresis* 2010, **31**, 2697-2704.
- [84] Jacobs, J. M., Adkins, J. N., Qian, W. J., Liu, T., *et al.*, Utilizing human blood plasma for proteomic biomarker discovery. *J Proteome Res* 2005, **4**, 1073-1085.
- [85] Gundry, R. L., White, M. Y., Nogee, J., Tchernyshyov, I., Van Eyk, J. E., Assessment of albumin removal from an immunoaffinity spin column: critical implications for proteomic examination of the albuminome and albumin-depleted samples. *Proteomics* 2009, **9**, 2021-2028.
- [86] Bellei, E., Bergamini, S., Monari, E., Fantoni, L. I., *et al.*, High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids* 2011, **40**, 145-156.
- [87] Capelo, J. L., Carreira, R., Diniz, M., Fernandes, L., *et al.*, Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques. *Anal Chim Acta* 2009, **650**, 151-159.
- [88] Hustoft, H. K., Reubsæet, L., Greibrokk, T., Lundanes, E., Malerød, H., Critical assessment of accelerating trypsination methods. *J Pharm Biomed Anal* 2011, **56**, 1069-1078.
- [89] Havlis, J., Thomas, H., Sebela, M., Shevchenko, A., Fast-response proteomics by accelerated in-gel digestion of proteins. *Anal Chem* 2003, **75**, 1300-1306.
- [90] Pramanik, B. N., Mirza, U. A., Ing, Y. H., Liu, Y. H., *et al.*, Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: a new approach to protein digestion in minutes. *Protein Sci* 2002, **11**, 2676-2687.
- [91] Lill, J. R., Ingle, E. S., Liu, P. S., Pham, V., Sandoval, W. N., Microwave-assisted proteomics. *Mass Spectrom Rev* 2007, **26**, 657-671.
- [92] Lin, S., Lin, Z., Yao, G., Deng, C., *et al.*, Development of microwave-assisted protein digestion based on trypsin-immobilized magnetic microspheres for highly efficient proteolysis followed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis. *Rapid Commun Mass Spectrom* 2007, **21**, 3910-3918.
- [93] Lin, S., Yao, G., Qi, D., Li, Y., *et al.*, Fast and efficient proteolysis by microwave-assisted protein digestion using trypsin-immobilized magnetic silica microspheres. *Anal Chem* 2008, **80**, 3655-3665.
- [94] Wang, S., Liu, T., Zhang, L., Chen, G., Efficient chymotryptic proteolysis enhanced by infrared radiation for peptide mapping. *J Proteome Res* 2008, **7**, 5049-5054.
- [95] Wang, S., Bao, H., Zhang, L., Yang, P., Chen, G., Infrared-assisted on-plate proteolysis for MALDI-TOF-MS peptide mapping. *Anal Chem* 2008, **80**, 5640-5647.
- [96] Wang, S., Zhang, L., Yang, P., Chen, G., Infrared-assisted tryptic proteolysis for peptide mapping. *Proteomics* 2008, **8**, 2579-2582.
- [97] Carreira, R. J., Lodeiro, C., Diniz, M. S., Moura, I., Capelo, J. L., Can ultrasonic energy efficiently speed (18)O-labeling of proteins? *Proteomics* 2009, **9**, 4974-4977.
- [98] Carreira, R. J., Lodeiro, C., Reboiro-Jato, M., Glez-Peña, D., *et al.*, Indirect ultrasonication for protein quantification and peptide mass mapping through mass spectrometry-based techniques. *Talanta* 2010, **82**, 587-593.
- [99] Fernández-Costa, C., Ruiz-Romero, C., Blanco, F. J., Santos, H. M., Capelo, J. L., An assessment of the indirect high intensity ultrasonic assisted cleavage of complex proteomes with immobilized trypsin. *Talanta* 2013, **106**, 163-168.

- [100] Ma, J., Zhang, L., Liang, Z., Zhang, W., Zhang, Y., Recent advances in immobilized enzymatic reactors and their applications in proteome analysis. *Anal Chim Acta* 2009, **632**, 1-8.
- [101] Krenkova, J., Lacher, N. A., Svec, F., Highly efficient enzyme reactors containing trypsin and endoproteinase LysC immobilized on porous polymer monolith coupled to MS suitable for analysis of antibodies. *Anal Chem* 2009, **81**, 2004-2012.
- [102] Krenkova, J., Svec, F., Less common applications of monoliths: IV. Recent developments in immobilized enzyme reactors for proteomics and biotechnology. *J Sep Sci* 2009, **32**, 706-718.
- [103] Yuan, H., Zhou, Y., Zhang, L., Liang, Z., Zhang, Y., Integrated protein analysis platform based on column switch recycling size exclusion chromatography, microenzymatic reactor and microRPLC-ESI-MS/MS. *J Chromatogr A* 2009, **1216**, 7478-7482.
- [104] Kim, J. H., Inerowicz, D., Hedrick, V., Regnier, F., Integrated sample preparation methodology for proteomics: analysis of native proteins. *Anal Chem* 2013, **85**, 8039-8045.
- [105] Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., *et al.*, From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y)* 1996, **14**, 61-65.
- [106] O'Farrell, P. H., High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 1975, **250**, 4007-4021.
- [107] Valcu, C. M., Valcu, M., Reproducibility of two-dimensional gel electrophoresis at different replication levels. *J Proteome Res* 2007, **6**, 4677-4683.
- [108] Unl , M., Morgan, M. E., Minden, J. S., Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 1997, **18**, 2071-2077.
- [109] Yates, J. R., Ruse, C. I., Nakorchevsky, A., Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng* 2009, **11**, 49-79.
- [110] Fournier, M. L., Gilmore, J. M., Martin-Brown, S. A., Washburn, M. P., Multidimensional separations-based shotgun proteomics. *Chem Rev* 2007, **107**, 3654-3686.
- [111] Righetti, P. G., Castagna, A., Antonioli, P., Boschetti, E., Prefractionation techniques in proteome analysis: the mining tools of the third millennium. *Electrophoresis* 2005, **26**, 297-319.
- [112] Motoyama, A., Yates, J. R., Multidimensional LC separations in shotgun proteomics. *Anal Chem* 2008, **80**, 7187-7193.
- [113] Issaq, H. J., Chan, K. C., Janini, G. M., Conrads, T. P., Veenstra, T. D., Multidimensional separation of peptides for effective proteomic analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, **817**, 35-47.
- [114] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* 1989, **246**, 64-71.
- [115] Tanaka, K., Waki, H., Ido, Y., Akita, S., *et al.*, Protein and polymer analysis up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectro.* 1988, **2**(8).
- [116] Karas, M., Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 1988, **60**, 2299-2301.
- [117] Blackstock, W. P., Weir, M. P., Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol* 1999, **17**, 121-127.

- [118] Pappin, D. J., Hojrup, P., Bleasby, A. J., Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol* 1993, 3, 327-332.
- [119] James, P., Quadroni, M., Carafoli, E., Gonnet, G., Protein identification by mass profile fingerprinting. *Biochem Biophys Res Commun* 1993, 195, 58-64.
- [120] Yates, J. R., Speicher, S., Griffin, P. R., Hunkapiller, T., Peptide mass maps: a highly informative approach to protein identification. *Anal Biochem* 1993, 214, 397-408.
- [121] Warder, S. E., Tucker, L. A., Strelitzer, T. J., McKeegan, E. M., *et al.*, Reducing agent-mediated precipitation of high-abundance plasma proteins. *Anal Biochem* 2009, 387, 184-193.
- [122] Kay, R., Barton, C., Ratcliffe, L., Matharoo-Ball, B., *et al.*, Enrichment of low molecular weight serum proteins using acetonitrile precipitation for mass spectrometry based proteomic analysis. *Rapid Commun Mass Spectrom* 2008, 22, 3255-3260.
- [123] Santos, H. M., Rial-Otero, R., Fernandes, L., Vale, G., *et al.*, Improving sample treatment for in-solution protein identification by peptide mass fingerprint using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J Proteome Res* 2007, 6, 3393-3399.
- [124] Gekko, K., Ohmae, E., Kameyama, K., Takagi, T., Acetonitrile-protein interactions: amino acid solubility and preferential solvation. *Biochim Biophys Acta* 1998, 1387, 195-205.
- [125] Oliviero, F., Sfriso, P., Baldo, G., Dayer, J. M., *et al.*, Apolipoprotein A-I and cholesterol in synovial fluid of patients with rheumatoid arthritis, psoriatic arthritis and osteoarthritis. *Clin Exp Rheumatol* 2009, 27, 79-83.
- [126] Takinami, Y., Yoshimatsu, S., Uchiumi, T., Toyosaki-Maeda, T., *et al.*, Identification of potential prognostic markers for knee osteoarthritis by serum proteomic analysis. *Biomark Insights* 2013, 8, 85-95.
- [127] Holme, I., Aastveit, A. H., Hammar, N., Jungner, I., Walldius, G., Inflammatory markers, lipoprotein components and risk of major cardiovascular events in 65,005 men and women in the Apolipoprotein MORTality RiSk study (AMORIS). *Atherosclerosis* 2010, 213, 299-305.
- [128] Miller, I., Goldfarb, M., in: Smejkal, G. B., Lazarev, A. (Eds.), *Separation Methods in Proteomics*, Taylor and Francis, London, UK. 2006.
- [129] Ruiz-Romero, C., Blanco, F. J., Achievements and challenges of proteomics in the study of rheumatoid arthritis. *Curr Top Med Chem* 2013, 13, 732-742.
- [130] Chang, X., Cui, Y., Zong, M., Zhao, Y., *et al.*, Identification of proteins with increased expression in rheumatoid arthritis synovial tissues. *J Rheumatol* 2009, 36, 872-880.
- [131] Ruiz-Romero, C., Blanco, F. J., Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis Cartilage* 2010, 18, 500-509.
- [132] Fernández, C., Santos, H. M., Ruiz-Romero, C., Blanco, F. J., Capelo-Martínez, J. L., A comparison of depletion versus equalization for reducing high-abundance proteins in human serum. *Electrophoresis* 2011, 32, 2966-2974.
- [133] Nedelkov, D., Kiernan, U. A., Niederkofler, E. E., Tubbs, K. A., Nelson, R. W., Investigating diversity in human plasma proteins. *Proc Natl Acad Sci U S A* 2005, 102, 10852-10857.
- [134] Alsaif, M., Guest, P. C., Schwarz, E., Reif, A., *et al.*, Analysis of serum and plasma identifies differences in molecular coverage, measurement variability, and candidate biomarker selection. *Proteomics Clin Appl* 2012, 6, 297-303.

- [135] Marouga, R., David, S., Hawkins, E., The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal Bioanal Chem* 2005, **382**, 669-678.
- [136] Lilley, K. S., Friedman, D. B., All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Rev Proteomics* 2004, **1**, 401-409.
- [137] Minden, J., Comparative proteomics and difference gel electrophoresis. *Biotechniques* 2007, **43**, 739, 741, 743 passim.
- [138] Olsen, A. K., Sondergaard, B. C., Byrjalsen, I., Tanko, L. B., *et al.*, Anabolic and catabolic function of chondrocyte ex vivo is reflected by the metabolic processing of type II collagen. *Osteoarthritis Cartilage* 2007, **15**, 335-342.
- [139] Wang, Q., Rozelle, A. L., Lepus, C. M., Scanzello, C. R., *et al.*, Identification of a central role for complement in osteoarthritis. *Nat Med* 2011, **17**, 1674-1679.
- [140] Blanco, F. J., Ruiz-Romero, C., Osteoarthritis: Metabolomic characterization of metabolic phenotypes in OA. *Nat Rev Rheumatol* 2012, **8**, 130-132.
- [141] Zhang, H., Li, X. J., Martin, D. B., Aebersold, R., Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 2003, **21**, 660-666.
- [142] Sinz, A., Bantscheff, M., Mikkat, S., Ringel, B., *et al.*, Mass spectrometric proteome analyses of synovial fluids and plasmas from patients suffering from rheumatoid arthritis and comparison to reactive arthritis or osteoarthritis. *Electrophoresis* 2002, **23**, 3445-3456.
- [143] Yamagiwa, H., Sarkar, G., Charlesworth, M. C., McCormick, D. J., Bolander, M. E., Two-dimensional gel electrophoresis of synovial fluid: method for detecting candidate protein markers for osteoarthritis. *J Orthop Sci* 2003, **8**, 482-490.
- [144] Gutteridge, J. M., The antioxidant activity of haptoglobin towards haemoglobin-stimulated lipid peroxidation. *Biochim Biophys Acta* 1987, **917**, 219-223.
- [145] Jue, D. M., Shim, B. S., Kang, Y. S., Inhibition of prostaglandin synthase activity of sheep seminal vesicular gland by human serum haptoglobin. *Mol Cell Biochem* 1983, **51**, 141-147.
- [146] Arredouani, M., Matthijs, P., Van Hoeyveld, E., Kasran, A., *et al.*, Haptoglobin directly affects T cells and suppresses T helper cell type 2 cytokine release. *Immunology* 2003, **108**, 144-151.
- [147] Capelo, J. L., Carreira, R. J., Fernandes, L., Lodeiro, C., *et al.*, Latest developments in sample treatment for 18O-isotopic labeling for proteomics mass spectrometry-based approaches: a critical review. *Talanta* 2010, **80**, 1476-1486.
- [148] Santos, H. M., Glez-Peña, D., Reboiro-Jato, M., Fdez-Riverola, F., *et al.*, A novel (18)O inverse labeling-based workflow for accurate bottom-up mass spectrometry quantification of proteins separated by gel electrophoresis. *Electrophoresis* 2010, **31**, 3407-3419.
- [149] Santos, H. M., Reboiro-Jato, M., Glez-Peña, D., Nunes-Miranda, J. D., *et al.*, Decision peptide-driven: a free software tool for accurate protein quantification using gel electrophoresis and matrix assisted laser desorption ionization time of flight mass spectrometry. *Talanta* 2010, **82**, 1412-1420.
- [150] Vale, G., Santos, H. M., Carreira, R. J., Fonseca, L., *et al.*, An assessment of the ultrasonic probe-based enhancement of protein cleavage with immobilized trypsin. *Proteomics* 2011, **11**, 3866-3876.
- [151] Fernández-Costa, C., Calamia, V., Fernández-Puente, P., Capelo-Martínez, J. L., *et al.*, Sequential depletion of human serum for the search of osteoarthritis biomarkers. *Proteome Sci* 2012, **10**, 55.

- [152] Tiss, A., Smith, C., Camuzeaux, S., Kabir, M., *et al.*, Serum peptide profiling using MALDI mass spectrometry: avoiding the pitfalls of coated magnetic beads using well-established ZipTip technology. *Proteomics* 2007, 7 Suppl 1, 77-89.
- [153] Callesen, A. K., Madsen, J. S., Vach, W., Kruse, T. A., *et al.*, Serum protein profiling by solid phase extraction and mass spectrometry: a future diagnostics tool? *Proteomics* 2009, 9, 1428-1441.
- [154] Bladergroen, M. R., Derks, R. J., Nicolardi, S., de Visser, B., *et al.*, Standardized and automated solid-phase extraction procedures for high-throughput proteomics of body fluids. *J Proteomics* 2012, 77, 144-153.
- [155] Mobasheri, A., Applications of proteomics to osteoarthritis, a musculoskeletal disease characterized by aging. *Front Physiol* 2011, 2, 108.
- [156] Zhang, A. H., Sun, H., Yan, G. L., Han, Y., Wang, X. J., Serum proteomics in biomedical research: a systematic review. *Appl Biochem Biotechnol* 2013, 170, 774-786.
- [157] Tjalsma, H., Identification of biomarkers for colorectal cancer through proteomics-based approaches. *Expert Rev Proteomics* 2010, 7, 879-895.
- [158] Böhm, D., Keller, K., Wehrwein, N., Lebrecht, A., *et al.*, Serum proteome profiling of primary breast cancer indicates a specific biomarker profile. *Oncol Rep* 2011, 26, 1051-1056.
- [159] Uto, H., Kanmura, S., Takami, Y., Tsubouchi, H., Clinical proteomics for liver disease: a promising approach for discovery of novel biomarkers. *Proteome Sci* 2010, 8, 70.
- [160] Roda, G., Caponi, A., Benevento, M., Nanni, P., *et al.*, New proteomic approaches for biomarker discovery in inflammatory bowel disease. *Inflamm Bowel Dis* 2010, 16, 1239-1246.
- [161] Park, J., Yang, J. S., Jung, G., Woo, H. I., *et al.*, Subunit-specific mass spectrometry method identifies haptoglobin subunit alpha as a diagnostic marker in non-small cell lung cancer. *J Proteomics* 2013.
- [162] Kamada, Y., Akita, M., Takeda, Y., Yamada, S., *et al.*, Serum Fucosylated Haptoglobin as a Novel Diagnostic Biomarker for Predicting Hepatocyte Ballooning and Nonalcoholic Steatohepatitis. *PLoS One* 2013, 8, e66328.
- [163] Orenes-Piñero, E., Hernández-Romero, D., de Torre, C., Vilchez, J. A., *et al.*, Identification and confirmation of haptoglobin as a potential serum biomarker in hypertrophic cardiomyopathy using proteomic approaches. *Ann Med* 2013, 45, 341-347.
- [164] Váradi, C., Mittermayr, S., Szekrényes, Á., Kádas, J., *et al.*, Analysis of haptoglobin N-glycome alterations in inflammatory and malignant lung diseases by capillary electrophoresis. *Electrophoresis* 2013, 34, 2287-2294.
- [165] Adams, J. N., Cox, A. J., Freedman, B. I., Langefeld, C. D., *et al.*, Genetic analysis of haptoglobin polymorphisms with cardiovascular disease and type 2 diabetes in the Diabetes Heart Study. *Cardiovasc Diabetol* 2013, 12, 31.
- [166] Chang, K. H., Tseng, M. Y., Ro, L. S., Lyu, R. K., *et al.*, Analyses of haptoglobin level in the cerebrospinal fluid and serum of patients with neuromyelitis optica and multiple sclerosis. *Clin Chim Acta* 2013, 417, 26-30.

7. ANEXOS

7.1. Anexo I: Material suplementario del artículo I

Table 1SM. List of 20 most abundant proteins in plasma.

Top 20	kDa
Albumin	71
IgG	36-40
IgA	38
IgM	50
IgD	42
Transferrin	79
Fibrinogen	50
Alpha-2-Macroglobulin	165
Alpha-1-Antitrypsin	47
Haptoglobin	46
Alpha-1-Acid Glycoprotein	24
Ceruloplasmin	123
Apolipoprotein A-I	31
Apolipoprotein A-II	11
Apolipoprotein B	517
Complement C1q	26
Complement C3	188
Complement C4	192
Plasminogen	93
Prealbumin	15

Table 2 SM. List of identified proteins.

Accession number	Protein's name	MW/pI	Peptide match	% Sequence coverage	Sequences	Score
P02763	Alpha-1-acid glycoprotein 1	23725/ 4.93	3	13	R.YVGGQEHFAHLILRLD R.YVGGQEHFAHLILRLDTK.T K.SDVVYTDWKK.D	139
P01009	Alpha-1-antitrypsin	46878/5.37	3	11	K.TDTSHHDDHPFNK.I K.ITPNLAEFASLYR.Q K.VFSNGADLSGVTEEAPLKLSK.A	48
P02750	Leucine-rich alpha-2-glycoprotein	38382/6.45	9	24	K.LQELHLSNGLESLSPEFLRPVPQLR.V K.ALGHLDLSGNRL K.ALGHLDLSGNRLR.K R.TLDLGENQLETLPDILLR.G R.TLDLGENQLETLPDILLRGLQLER.L R.GPLQLER.L R.YLFLNGNK.L R.YLFLNGNKLAR.V R.VAAGAFQGLR.Q	209
P01023	Alpha-2-macroglobulin	164614/6.00	14	10	K.SGGRTEHPFTVEEFVLPK.F K.VDSHERQGIFFGQVRL R.QGIPFFGQVRL K.GVPIPNKVIFIR.G K.VDLSFSPQSPLPASHAHLR.V R.VTAAPOSVCALR.A R.VGFYESDVMGRGHAR.L R.LVHVVEPHITETVR.K K.SKAIGYLNITGYQR.Q K.AIGYLNITGYQR.Q R.QLNYKHVDGSYSITFGER.Y K.KDNSVHWERPQKPK.A K.DNSVHWERPQKPK.A K.YGAATFTR.T	509
P01019	Angiotensinogen	53406/5.87	2	7	K.ANAGKPKDPTFIPAPIQAK.T K.QPFVQGLALYTPVVLPR.S	131
P01008	Antithrombin-III	53025/6.32	4	10	R.DIPMNPMTIYR.S K.ATEDEGSEQIPEATNR.R R.RVWELSK.A K.LQPLDFKENAEQSR.A	236
P02647	Apolipoprotein A-I	30759/5.56	8	32	K.DSGRDYVSQFEGSALGK.Q K.LLDNWDSTSTFSK.L K.VEPLRAELQEGAR.Q R.THLAPYSDELRL R.THLAPYSDELRLR.L R.LEALKENGGAR.L K.ATEHLSTLSEKAKPALEDLR.Q K.KPALEDLR.Q	226
P02652	Apolipoprotein A-II	11282 /6.26	2	11	K.VKSPQLQAEAK.S K.SPQLQAEAK.S	89
P06727	Apolipoprotein A-IV	45371/5.28	9	23	K.KLVPFATELHER.L K.LVPFATELHER.L R.LEPYADQLR.T R.QLTPYAQR.M R.ISASAEELRQR.L R.LAPLAEDVRGNLR.G K.SLAELGGHLDQQVEEFRR.R R.RVEPYGENFNK.A K.ALVOQMEQLR.Q	126
P04114	Apolipoprotein B-100	516666/6.58	7	2	K.SVGHILPSREFQVPTFTIPK.L R.DLKVEDIPLAR.I R.VPSYTLILPSLELPVLHVPR.N K.FSVPAIVIPISFQALTAR.F K.SPAFTDLHLR.Y R.NLQNNAEWVYQGAIK.Q K.IISDYHQFRR.Y	115
P02656	Apolipoprotein C-III	10843/5.23	2	19	K.TAKDALSSVQESQVAQAR.G K.DALSSVQESQVAQAR.G	252
P05090	Apolipoprotein D	21547/5.06	5	22	K.CPNPPVQENFDVNYGLRW R.WYEIEKIPTTFENGR.C K.IPTTFENGR.C K.IKVLNQELR.A K.VLNQELR.A	229
P02649	Apolipoprotein E	36246/5.65	8	23	R.FWDYLR.W K.AYKSELEEQLTPVAEETRA R.LAVYQAGARE R.LGPLVEQGR.V R.AATVGSLAGQPLQER.A R.AQAWGER.L R.AQAWGERLR.A	251
O95445	Apolipoprotein M	21582/5.66	5	22	R.KWIYHLEGTSTDLR R.FLLYNR.S R.FLLYNRSPHPPEK.C K.SLTSCLDSKAFLLTPR.N K.AFLLLTPR.N	95
P00450	Ceruloplasmin	122983/ 5.44	6	5	K.ALYLYQTDFTFR.T R.VTFHNKGAYPLSIEPIGVRF K.GAYPLSIEPIGVRF R.FNKNNEGTYYSPPNYNPSQR.S K.NNEGTYYSPPNYNPSQR.S R.QYTDSTFRVPVER.K	216
P10909	Clusterin	53031/5.89	6	18	R.ASSIIDELFQDRFFTR.E R.EPQDTHYLPFSLPHR.R R.RPHFFFPK.S R.ELDESQVAER.L K.LFSDSPITVTPVEVSR.K K.FMETVAEKALQEYR.K	302

P01024	Complement C3	188569/6.02	11	7	R.EGVQKEDIPPADLSDQVPDTESETR.I R.ILLQGTTPVAQMTEDAVDAERL.L R.ILLQGTTPVAQMTEDAVDAERL.K.H K.KGYTQQLAFR.Q K.GYTQQLAFR.Q K.GYTQQLAFRQPSAFAAFVK.R R.QPSSAFAAFVKR.A K.QKPDGVFQEDAPVIHQEMIGGLR.N K.AGDFLEANYMNLQR.S R.SYTVAIAGYALAQMGRL R.WLNEQR.Y	592
P0C0L4	Complement C4-A	194247/6.65	10	3	R.RCSVFYGAPSK.S R.CSVFYGAPSK.S R.CSVFYGAPSKSR.L R.LLATLCSAEVCQCAEGK.C R.LLATLCSAEVCQCAEGKCPR.Q R.GLQDEDDGYR.M R.GLQDEDDGYR.M.F K.FACYYP.R.V K.FACYYP.RVEYGFQVK.V K.ITQVLHFTK.D	469
P02748	Complement component C9	64615/5.43	2	4	R.SIEVFQFNGKR.C R.AIEDYINEFSVRK.C	130
P00751	Complement factor B	86847/6.67	2	3	R.AIHCPRPDHFENGEYWP.R.S R.YGLVTYATYPI	37
P08603	Complement factor H	143680/6.21	5	3	R.NGFYPATR.G R.NGFYPATRGNTAK.C K.CTSTGWIPAPR.C K.SCDIPVFMNAR.T R.QMSKYP.SGER.V	77
Q03591	Complement factor H-related protein 1	38766/7.38	3	9	K.INHGILYDEEK.Y R.QMSKYP.SGER.V R.EIMENYNIALR.W	171
P02751	Fibronectin	266034/ 5.46	12	7	R.TYLGNALVCTCYGGS.R.G R.YQCYCYGR.G R.WSRPQAPITGYR.I R.QYNVGPSVSKYPLR.N K.LGVRPSQGEAPR.E R.FTNIGPDTMR.V R.VTWAPPSIDLTNFLVR.Y	150
P23142	Fibulin-1	81268/5.07	3	4	R.CCHCCLLR.A K.TGYFDGISR.M R.CLAFECPENYRR.S	66
Q08380	Galectin-3-binding protein	66202/ 5.13	4	9	K.AVDTSWGER.A K.TLQALEFHTVPFQLLAR.Y K.YSSDYFQAPSDYR.Y K.TSSSFPCAGHFNFR.T	161
P22352	Glutathione peroxidase 3	25765/8.26	5	25	K.YVVRPGGFVNPQLFEK.G K.FYTLK.N K.NSCPPTSELLGTSDR.L K.FLVGPDGIPMR.W K.MDILSYMR.R	88
P00738	Haptoglobin	45861/6.13	9	23	R.ILGHLDAKGSFPWQAK.M K.GSFPWQAK.M R.VMPICLPKDYAEVGR.V R.VMPICLPKDYAEVGR.V R.VGYVSGWGR.N R.NANFKFTDHLK.Y K.YVVMLPVADQDQCIR.H K.SPVGVQPILNEHTFCAGMSK.Y K.VTSIQDWVQK.T	399
P69905	Hemoglobin subunit alpha	15305/8.72	3	28	K.VGAHAGEYGAELER.M R.MFLSFPTTK.T K.TYFPHFDLSHGSAQVK.G	306
P68871	Hemoglobin subunit beta	16102/6.75	3	25	K.VNVDEVGGEALGR.L R.LLVVYPWTQRF K.VVAGVANALAHKYH.-	228
P01876	Ig alpha-1 chain C region	38486/6.08	7	19	R.DASGVTFWTWTPSSGK.S K.SAVQGPPER.D R.GFSPKDV.LVR.W R.WLQGSQELPRE.E R.WLQGSQELPREK.Y K.YLTWASR.Q R.QEPSOGTTTFAVTSILR.V	323
P01877	Ig alpha-2 chain C region	37301/ 5.71	4	12	R.GFSPKDV.LVR.W R.WLQGSQELPRE.E K.YLTWASR.Q R.QEPSOGTTTFAVTSILR.V	183
P01857	Ig gamma-1 chain C region	36596/ 8.46	4	13	K.GPSVFPLAPSSK.S K.ALPAPIEK.T K.ALPAPIEKTIK.A K.GFYPSDIAVWESNGQPENNYK.T	146
P01859	Ig gamma-2 chain C region	36505/7.66	4	12	K.GPSVFPLAPCSR.S K.VSNKGLPAPIEK.T K.GLPAPIEK.T R.EPQVYTLPPSREEMTK.N	69
P01860	Ig gamma-3 chain C region	42287/8.23	4	10	K.GPSVFPLAPCSR.S K.ALPAPIEK.T K.ALPAPIEKTIK.T R.EPQVYTLPPSREEMTK.N	88
P01861	Ig gamma-4 chain C region	36431/ 7.18	2	10	K.GPSVFPLAPCSR.S K.GFYPSDIAVWESNGQPENNYK.T	108
P01766	Ig heavy chain V-III region BRO	13332/ 6.45	1	15	-EVQLVESGGGLVQPGGSLR.L	147
P01834	Ig kappa chain C region	11773/ 5.58	4	41	-TVAAPSVPFIPPSDEQLK.S K.SGTASVVCLLNIFYPRE.E K.SGTASVVCLLNIFYPREAK.V K.SFNREGC.-	216

P01620	Ig kappa chain V-III region SIE	11882/ 8.70	1	16	-EIVLTQSPGTLSPGER.A	97
P01625	Ig kappa chain V-IV region Len	12746/ 7.92	1	15	-DIVMTQSPDSLAVSLGER.A	85
P0CG05	Ig lambda chain C regions	11401/ 6.92	3	37	K.AAPSVTLFPPSSEELQANK.A K.QSNKKYAASSYLSLTPEQWK.S K.YAASSYLSLTPEQWK.S	126
P01717	Ig lambda chain V-IV region Hil	11624/ 6.04	1	17	-SYELTQPPSVSVSPGQTAR.I	87
P0CG05	Ig lambda-2 chain C regions	11458/ 6.92	3	36	K.AAPSVTLFPPSSEELQANK.A K.QSNKKYAASSYLSLTPEQWK.S K.YAASSYLSLTPEQWK.S	120
P01871	Ig mu chain C region	49960/ 6.35	5	11	K.SKLCQATGFSR.Q K.LICQATGFSR.Q R.QIQVSWLR.E K.GVALHRPDVYLLPPAR.E K.YVTSAPMPEQAPGR.Y	110
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	101782/ 6.31	3	5	R.FAHYVVTSQVNTANEAR.E K.VTFQLTYEEVLKR.N R.GFSLDEATNLNGGLR.G	40
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	106853/ 6.40	3	3	K.AHVSFKPTVAQQR.I R.LSNENHGIAQR.I K.FYNQVSTPLLR.N	136
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	103521/ 6.51	2	1	R.FAHTVVTSR.V K.VRPQQLVK.H	43
P02750	Leucine-rich alpha-2-glycoprotein	38382/ 6.45	9	24	K.LQELHLSNGLESLSPEFLRPVPQLR.V K.ALGHLDL.SGNNR.L K.ALGHLDL.SGNNR.L.K R.TLDLGENQLETLPPDLLR.G R.TLDLGENQLETLPPDLLRGPLQLER.L R.GPLQLER.L R.YLFLNGNK.L R.YLFLNGNKLAR.V R.VAAGAFQGLR.Q	209
Q12912	Lymphoid-restricted membrane protein	62753/ 5.62	1	1	R.VTIASLPR.N	32
P05155	Plasma protease C1 inhibitor	55347/ 6.09	2	4	K.KYPVAHFIDQTLK.A K.FQPTLLTLPR.I	43
P00747	Plasminogen	93247/ 7.04	1	1	R.HSIFTPETNPR.A	51
P00734	Prothrombin	71475/ 5.64	3	8	R.SGIECQLWR.S R.SEGSSVNLSPPLEQCVPRGQQYQGR.L R.LAVTTHGLPCLAWASQAQ.A	107
P02753	Retinol-binding protein 4	23337/ 5.76	5	22	R.FSGTWYAMAK.K	212
P02787	Serotransferrin	79280/ 6.81	4	8	K.DSGFQMNQLR.G R.SAGWNPIGLLYCDLPEPR.K K.MYLGVEYVTAIR.N K.NLNEKDYELLCLDGTR.K	121
P02768	Serum albumin	71317/ 5.92	14	26	K.SLHTLFGDKLCTVATLR.E K.YLYEIAIR.R K.YLYEIAIR.H R.RHPYFYAPELFFAK.R R.AFKAWAVAR.L R.LSORFPK.A K.VHTECCHGDLLCEADDRADLAK.Y K.NYAEAKDVFLGMFLYEYAR.R R.RHPDYSVVLLRL.L R.HPDYSVVLLRL.L K.VFDEFKPLVEEPQNLK.Q K.QNCELFEOLGEYKFQNALVR.Y K.KVPQVSTPTLVEVSR.N K.VPQVSTPTLVEVSR.N	186
P02735	Serum amyloid A protein	13581/ 6.28	2	27	R.EANYIGSDKYFHAR.G R.FFGHGAEDSLADQAAANEWGR.S	148
P35542	Serum amyloid A-4 protein	14854/ 9.17	4	43	R.SFFKEALQGVGDMGR.A R.AYWDIMISNHQNSNR.Y R.GNYDAAQRGPGGVWAAK.L K.SNEKAEWGR.S	138
P02743	Serum amyloid P-component	25485/ 6.10	5	22	R.AYSLSFYNTQGR.D R.VGEYSLYIGR.H K.GLROGYFVEAQPK.I R.OGYFVEAQPK.I K.IVLGOEQDSYGGKFDR.S	209
P02766	Transthyretin	15991/ 5.52	6	34	R.GSPAINVAVHVER.K R.GSPAINVAVHVER.A R.KAADDTWEPFASGK.T K.AADDTWEPFASGK.T K.ALGISPFHEHAIEVVTANDSGPR.R K.ALGISPFHEHAIEVVTANDSGPRR.Y	377
P04004	Vitronectin	55069/ 5.55	10	20	R.GQVCYELDEKAVRPGYPK.L K.AVRPGYPK.L K.LIRDVWGIEGPIDAAAFTR.I R.DVWGIEGPIDAAAFTR.I R.FEDGVLDPDYPR.N	362
Q86VM9	Zinc finger CCCHdomain-containing protein 18	106544/ 8.39	1	1	K.KLGVSVSPSR.A	52
P25311	Zinc-alpha-2-glycoprotein	34465/ 5.71	2	7	K.AREDIFMETLK.D K.QKWEAEPVYVQRA	45

Figure 1SM. Comprehensive scheme of the workflow followed in this research to assess the three methods for abundant protein reduction chosen for this study. The workflow was designed to take into account the technical variability of the experiments.

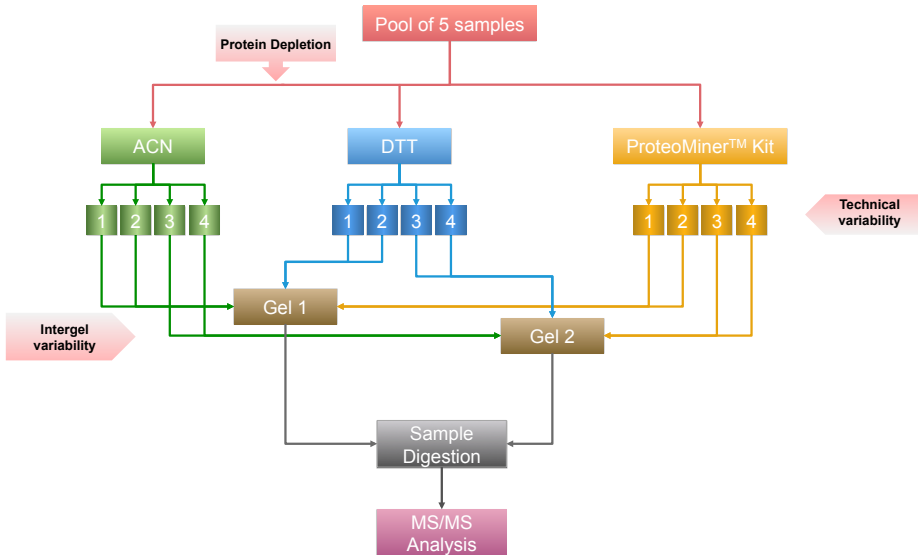


Figure 2SM. Bands excised and proteins identified in each band for two replicates for non-depleted serum. Each replicate means and independent band. Lines 1 and 2 were run in the same gel.

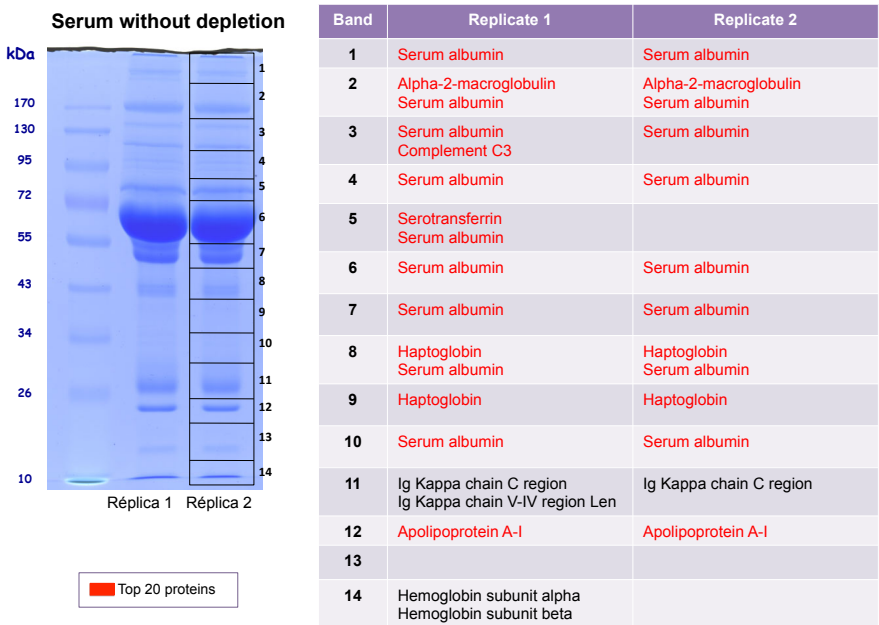


Figure 3SM. Gels done to develop this work.

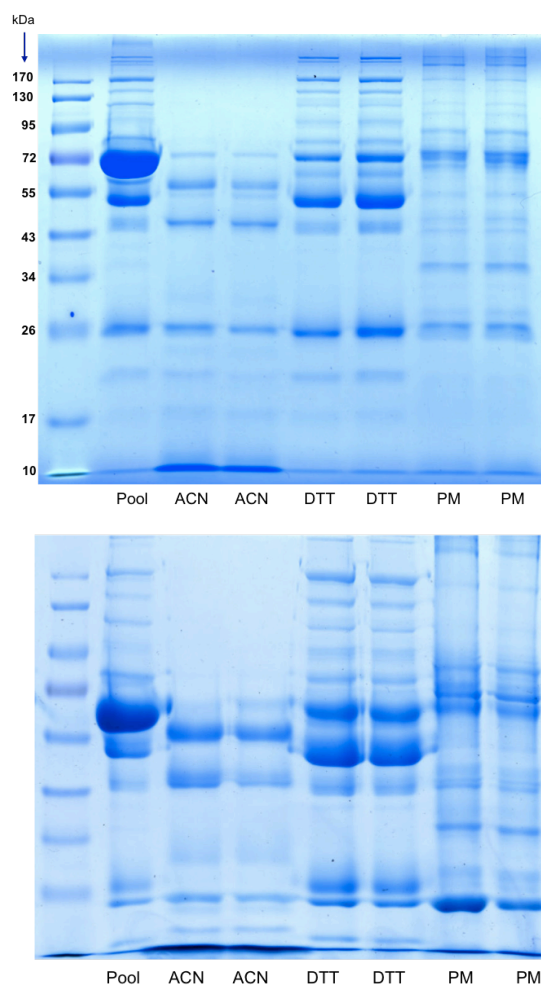


Figure 4SM. Bands excised and proteins identified in each band for four replicates after protein depletion with 57% V/V ACN. Each replicate means and independent band. Lines 1 and 2 were run in the same gel.

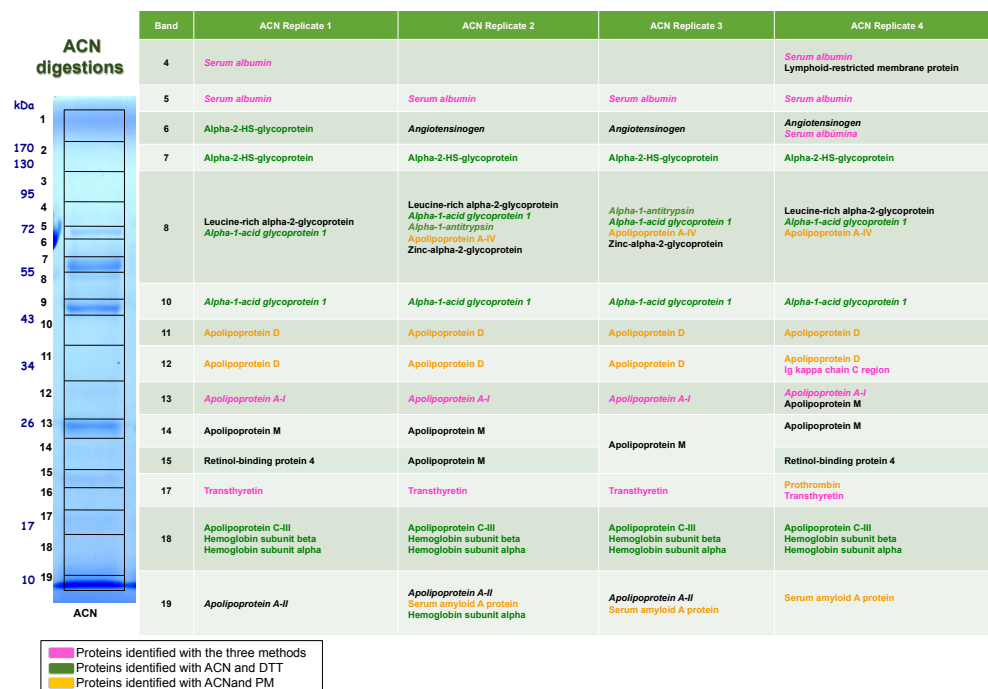


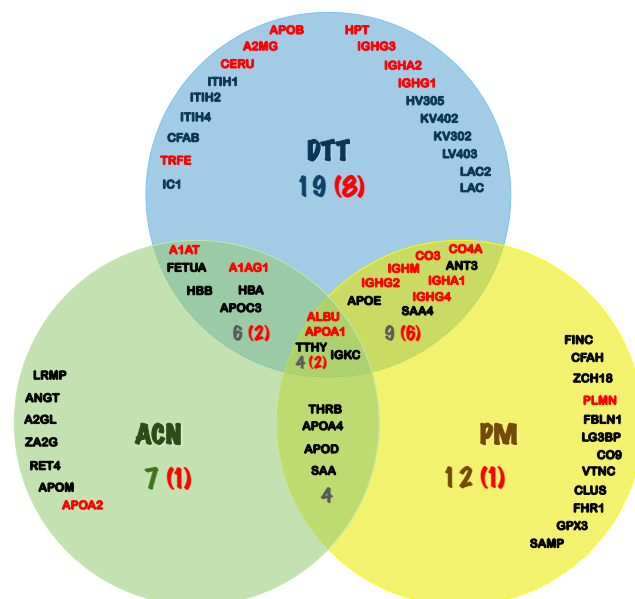
Figure 5SM. Bands excised and proteins identified in each band for four replicates after protein depletion with 500 mM DTT. Each replicate means and independent band. Lines 1 and 2 were run in the same gel.



Figure 6SM. Bands excised and proteins identified in each band for four replicates after protein depletion with PM. Each replicate means and independent band. Lines 1 and 2 were run in the same gel.



Figure 7SM. Venn diagrams showing the proteins identified with the three methods studied in this work. Red color: proteins belonging to the 20 most abundant proteins. Black color: Proteins out of the top 20.



Protein abbreviations

A1AG1: Alpha-1-acid glycoprotein 1; A1AT: Alpha-1-antitrypsin; A2GL: Leucine-rich alpha-2-glycoprotein; A2MG: Alpha-2-macroglobulin; ALBU: Serum albumin; ANG1: Angiotensinogen; ANT3: Antithrombin-III; APOA1: Apolipoprotein A-I; APOA2: Apolipoprotein A-II; APOA4: Apolipoprotein A-IV; APOB: Apolipoprotein B-100; APOC3: Apolipoprotein C-III; APOD: Apolipoprotein D; APOE: Apolipoprotein E; APOM: Apolipoprotein M; CERU: Ceruloplasmin; CFAB: Complement factor B; CFAH: Complement factor H; CLUS: Clusterin; CO3: Complement C3; CO4A: Complement C4-A; CO9: Complement component C9; FBLN1: Fibulin-1; FETUA: Alpha-2-HS-glycoprotein; FHR1: Complement factor H-related protein 1; FINC: Fibronectin; GPX3: Glutathione peroxidase 3; HBA: Hemoglobin subunit alpha; HBB: Hemoglobin subunit beta; HPT: Haptoglobin; HV305: Ig heavy chain V-III region BRO; IC1: Plasma protease C1 inhibitor; IGHA1: Ig alpha-1 chain C region; IGHA2: Ig alpha-2 chain C region; IGHG1: Ig gamma-1 chain C region; IGHG2: Ig gamma-2 chain C region; IGHG3: Ig gamma-3 chain C region; IGHG4: Ig gamma-4 chain C region; IGHM: Ig mu chain C region; IGKC: Ig kappa chain C region; ITIH1: Inter-alpha-trypsin inhibitor heavy chain H1; ITIH2: Inter-alpha-trypsin inhibitor heavy chain H2; ITIH4: Inter-alpha-trypsin inhibitor heavy chain H4; KV302: Ig kappa chain V-III region SIE; KV402: Ig kappa chain V-IV region Len; LAC: Ig lambda chain C regions; LAC2: Ig lambda-2 chain C regions; LG3BP: Galectin-3-binding protein; LRMP: Lymphoid-restricted membrane protein; LV403: Ig lambda chain V-IV region Hil; PLMN: Plasminogen; RET4: Retinol-binding protein 4; SAA: Serum amyloid A protein; SAA4: Serum amyloid A-4 protein; SAMP: Serum amyloid P-component; THRB: Prothrombin; TRFE: Serotransferrin; TTHY: Transthyretin; VTNC: Vitronectin; ZA2G: Zinc-alpha-2-glycoprotein; ZCH18: Zinc finger CCCHdomain-containing protein 18.

7.2. Anexo II: Material suplementario del artículo II

Figure S1. Representative 2D map of the sequentially depleted serum, with the protein identifications that were carried out. Identifications are listed in Additional file 2: Table S1.

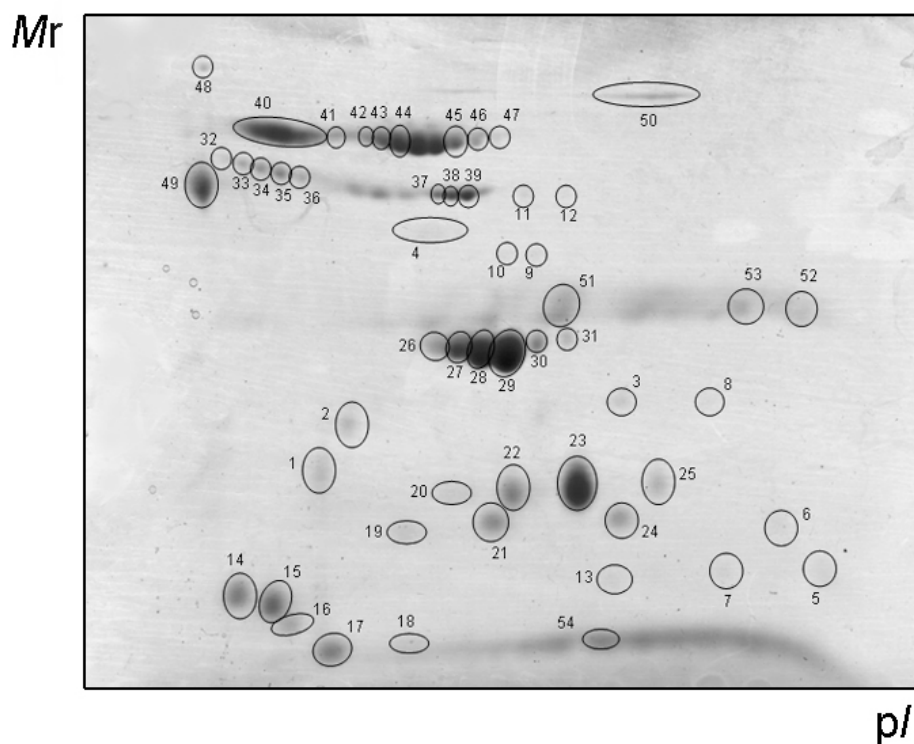


Table S1. Mass spectrometry data of the protein identifications performed in the present study. Spot numbers, according to Additional file 1: Figure S1.

Spot	Mascot identification	<u>Spots identified by MSMS</u>	Sec. Cov.	pI
3	HPT HUMAN Mass: 45861 Score: 135 Matches: 4(4) Sequences: 3(3) Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1		11%	6.13
5	HBB HUMAN Mass: 16102 Score: 43 Matches: 1(1) Sequences: 1(1) Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2		6%	6.75
6	HPTR HUMAN Mass: 39518 Score: 35 Matches: 1(1) Sequences: 1(1) Haptoglobin-related protein OS=Homo sapiens GN=HPR PE=2 SV=2		4%	6.63
7	SAA HUMAN Mass: 13581 Score: 52 Matches: 2(1) Sequences: 2(1) Serum amyloid A protein OS=Homo sapiens GN=SAA1 PE=1 SV=2		27%	6.28
8	TTHY HUMAN Mass: 15991 Score: 62 Matches: 2(1) Sequences: 2(1) Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1		9%	5.52
10	HPT HUMAN Mass: 45861 Score: 45 Matches: 5(2) Sequences: 4(2) Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1		13%	6.13
13	SAA HUMAN Mass: 13581 Score: 53 Matches: 4(2) Sequences: 4(2) Serum amyloid A protein OS=Homo sapiens GN=SAA1 PE=1 SV=2		52%	6.28
14	APOC3 HUMAN Mass: 10846 Score: 123 Matches: 2(2) Sequences: 2(2) Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1		19%	5.23
15	APOC3 HUMAN Mass: 10846 Score: 149 Matches: 2(2) Sequences: 2(2) Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1		19%	5.23
17	APOA2 HUMAN Mass: 11282 Score: 82 Matches: 3(2) Sequences: 3(2) Apolipoprotein A-II OS=Homo sapiens GN=APOA2 PE=1 SV=1		21%	6.26
20	TTHY HUMAN Mass: 15991 Score: 136 Matches: 5(3) Sequences: 5(3) Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1		34%	5.52
21	HPT HUMAN Mass: 45861 Score: 96 Matches: 3(2) Sequences: 3(2) Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1		9%	6.13
22	TTHY HUMAN Mass: 15991 Score: 205 Matches: 6(5) Sequences: 6(5) Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1		38%	5.52
23	TTHY HUMAN Mass: 15991 Score: 415 Matches: 9(7) Sequences: 9(7) Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1		73%	5.52
24	HPT HUMAN Mass: 45861 Score: 68 Matches: 3(2) Sequences: 3(2) Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1		9%	6.13
25	TTHY HUMAN Mass: 15991 Score: 57 Matches: 3(2) Sequences: 3(2) Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1		18%	5.52
26	APOA1 HUMAN Mass: 30759 Score: 97 Matches: 8(5) Sequences: 8(5) Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1		23%	5.56
27	APOA1 HUMAN Mass: 30759 Score: 125 Matches: 9(4) Sequences: 9(4) Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1		33%	5.56
28	APOA1 HUMAN Mass: 30759 Score: 271 Matches: 12(7) Sequences: 12(7) Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1		41%	5.56

29	APOA1 HUMAN Mass: 30759 Score: 305 Matches: 10(7) Sequences: 10(7) Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	36%	5.56
30	APOA1 HUMAN Mass: 30759 Score: 42 Matches: 5(2) Sequences: 5(2) Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	16%	5.56
31	APOA1 HUMAN Mass: 30759 Score: 135 Matches: 6(5) Sequences: 6(5) Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	20%	5.56
32	A2GL HUMAN Mass: 38382 Score: 40 Matches: 3(2) Sequences: 3(2) Leucine-rich alpha-2-glycoprotein OS=Homo sapiens GN=LRG1 PE=1 SV=2	11%	6.45
33	A2GL HUMAN Mass: 38382 Score: 88 Matches: 3(2) Sequences: 3(2) Leucine-rich alpha-2-glycoprotein OS=Homo sapiens GN=LRG1 PE=1 SV=2	10%	6.45
34	A2GL HUMAN Mass: 38382 Score: 36 Matches: 1(1) Sequences: 1(1) Leucine-rich alpha-2-glycoprotein OS=Homo sapiens GN=LRG1 PE=1 SV=2	2%	6.45
35	A2GL HUMAN Mass: 38382 Score: 29 Matches: 1(1) Sequences: 1(1) Leucine-rich alpha-2-glycoprotein OS=Homo sapiens GN=LRG1 PE=1 SV=2	2%	6.45
37			
38	APOA4 HUMAN Mass: 45371 Score: 128 Expect: 3.2e-09 Matches: 11 Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3	25%	5.28
39	APOA4 HUMAN Mass: 45371 Score: 37 Matches: 13(2) Sequences: 12(2) Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3	28%	5.28
40	FETUA HUMAN Mass: 40098 Score: 102 Matches: 3(2) Sequences: 2(2) Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	8%	5.43
41	FETUA HUMAN Mass: 40098 Score: 48 Matches: 1(1) Sequences: 1(1) Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	3%	5.43
42	A1AT HUMAN Mass: 46878 Score: 56 Matches: 6(4) Sequences: 6(4) Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	19%	5.37
43	A1AT HUMAN Mass: 46878 Score: 151 Matches: 10(5) Sequences: 10(5) Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	32%	5.37
44	A1AT HUMAN Mass: 46878 Score: 497 Matches: 14(11) Sequences: 13(10) Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	37%	5.37
45	A1AT HUMAN Mass: 46878 Score: 156 Matches: 11(7) Sequences: 10(7) Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	32%	5.37
46	ANGT HUMAN Mass: 53406 Score: 28 Matches: 1(1) Sequences: 1(1) Angiotensinogen OS=Homo sapiens GN=AGT PE=1 SV=1	3%	5.87
47	ANGT HUMAN Mass: 53406 Score: 26 Matches: 1(1) Sequences: 1(1) Angiotensinogen OS=Homo sapiens GN=AGT PE=1 SV=1	3%	5.87
49	A1AG1 HUMAN Mass: 23725 Score: 167 Matches: 6(5) Sequences: 6(5) Alpha-1-acid glycoprotein 1 OS=Homo sapiens GN=ORM1 PE=1 SV=1	22%	4.93
51	IGKC HUMAN Mass: 11773 Score: 106 Matches: 2(2) Sequences: 2(2) Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	32%	5.58
	LAC2 HUMAN Mass: 11458 Score: 54 Matches: 3(2) Sequences: 3(2) Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1	46%	6.92
52	IGKC HUMAN Mass: 11773 Score: 107 Matches: 2(2) Sequences: 2(2) Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	2%	5.58
53	LAC2 HUMAN Mass: 11458 Score: 72 Matches: 4(2) Sequences: 4(2) Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1	58%	6.92
	IGKC HUMAN Mass: 11773 Score: 50 Matches: 3(1) Sequences: 3(1) Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	34%	5.58

		Spots identified by MS		
Spot	Mascot identification		Sec. Cov.	pI
37	APOA4 HUMAN Mass: 45371 Score: 90 Expect: 2.1e-05 Matches: 8 Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3		20%	5.28
38	APOA4 HUMAN Mass: 45371 Score: 156 Expect: 5.1e-12 Matches: 13 Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3		29%	5.28
52	LAC2 HUMAN Mass: 11458 Score: 60 Expect: 0.019 Matches: 4 Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1		50%	6.92

Search Parameters

Type of search : MS/MS Ion Search
Enzyme : Trypsin
Fixed modifications : Carbamidomethyl (C)
Variable modifications : Oxidation (M)
Mass values : Monoisotopic
Protein Mass : Unrestricted
Peptide Mass Tolerance : ± 50 ppm
Fragment Mass Tolerance: ± 0.3 Da
Max Missed Cleavages : 1
Instrument type : MALDI-TOF-TOF
Number of queries : 15

Database : SwissProt 2011_05 (528048 sequences; 186939477 residues)
Taxonomy : Homo sapiens (human) (20239 sequences)
Timestamp : 7 Jun 2011 at 16:38:17 GMT

Figure S2. Amino acid sequence of Haptoglobin, with the MS data corresponding to the identification of alpha and beta chains.

P00738, Haptoglobin, Homo sapiens

10	20	30	40	50	60
MSALGAVIAL	LLWGQLFAVD	SGNDVTDIAD	DGCPKPPEIA	HGYVEHSVRV	QCKNYYKLRT
70	80	90	100	110	120
EGDGVYTLND	KKQWINKAVG	DKLPECEADD	GCPKPPEIAH	GYVEHSVRVQ	CKNYYKLRT
130	140	150	160	170	180
GDGVYTLNNE	KQWINKAVGD	KLPECEAVCG	KPKNPANPVQ	RILGGHLDK	GSFPWQAKMV
190	200	210	220	230	240
SHHNLTTGAT	LINEQWLLTT	AKNLFNLHSE	NATAKDIAPT	LTLYVGKKQL	VEIEKVVLHP
250	260	270	280	290	300
NYSQVDIGLI	KLKQKVSUNE	RVMPICLPSP	DYAEVGRVGY	VSGWGRNANF	KFTDHLKYVM
310	320	330	340	350	360
LPVADQDQCI	RHYEGSTVPE	KKTPKSPVGV	QPIILNEHTFC	AGMSKYQEDT	CYGDAGSAFA
370	380	390	400		
VHDLLEEDTWY	ATGILSFDKS	CAVAEYGVYV	KVTSIQDWVQ	KTIAEN	

Haptoglobin chains:

[162-406], Haptoglobin beta chain (yellow shadow), 27.27kDa

[19-160], Haptoglobin alfa chain, 15.95kDa

Peptides identified as Haptoglobin alfa chain (spot 3)

[HPT HUMAN](#) [Mass: 45861](#) [Score: 135](#) [Matches: 4\(4\)](#) [Sequences: 3\(3\)](#)

Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1

Start-End	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Sequence
60 - 72	1439.7435	1438.7362	1438.6940	29	1	R.TEGDGVYTLNDKK.Q (lons score 50)
117 - 131	1708.8474	1707.8401	1707.8428	-2	1	K.LRTEGDGVYTLNNEK.Q (lons score 28)
137 - 153	1857.9100	1856.9027	1856.9124	-5	1	K.AVGDKLPECEAVCGKPK.N (lons score 75)
137 - 153	1858.0114	1857.0041	1856.9124	49	1	K.AVGDKLPECEAVCGKPK.N (lons score 61)

Matched peptides shown in **Red**

Peptides identified as Haptoglobin beta chain (spot 10)

[HPT HUMAN](#) [Mass: 45861](#) [Score: 45](#) [Matches: 5\(2\)](#) [Sequences: 4\(2\)](#)

Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1

Start-End	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Sequence
171 - 178	920.4519	919.4446	919.4552	-12	0	K.GSFPWQAK.M (lons score 15)
287 - 297	1334.6849	1333.6776	1333.6779	-0	1	R.NANFKFTDHLK.Y (lons score 39)
298 - 311	1707.8236	1706.8163	1706.8120	3	0	K.YVMLPVADQDQCIR.H (lons score 32)
298 - 311	1723.8199	1722.8126	1722.8069	3	0	K.YVMLPVADQDQCIR.H Oxidation (M) (lons score 27)
326 - 345	2188.0701	2187.0628	2187.0453	8	0	K.SPVGVQPILNEHTFCAGMSK.Y Oxidation (M) (lons score 7)

Matched peptides shown in **Blue**

7.3. Anexo III: Material suplementario del artículo III

Figure 1SM. The peptide m/z 1465. This peptide has double ^{18}O incorporation efficiency higher than 95%. When the label and non-label counterpart are mixed, the spectra of both do not mix. B: The peptide m/z 1226, this peptide has a double incorporation of about 70%. When the label and the non-label counterpart are 1:1 mixed, the spectra of both mix, compromising quantification.

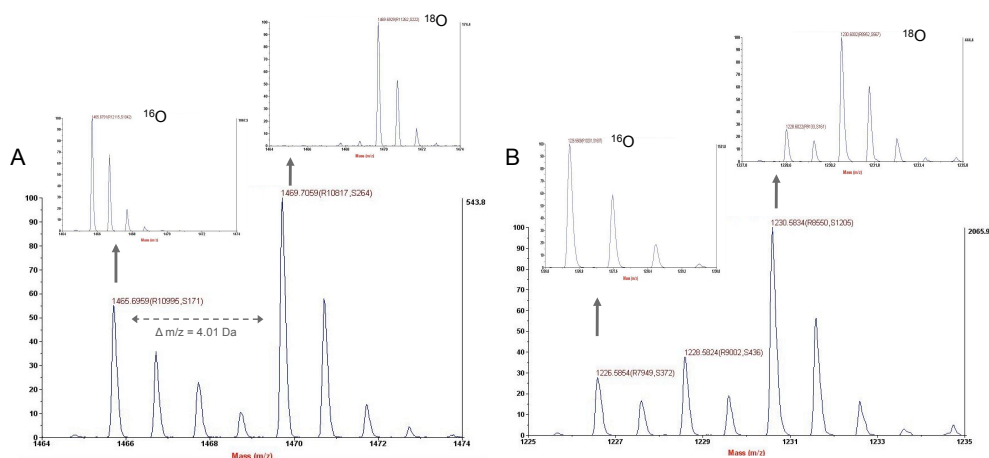


Figure 2SM. Comprehensive scheme of the sample treatment followed to assess the influence of the type of ultrasonic device used in a digestion done with immobilized trypsin.

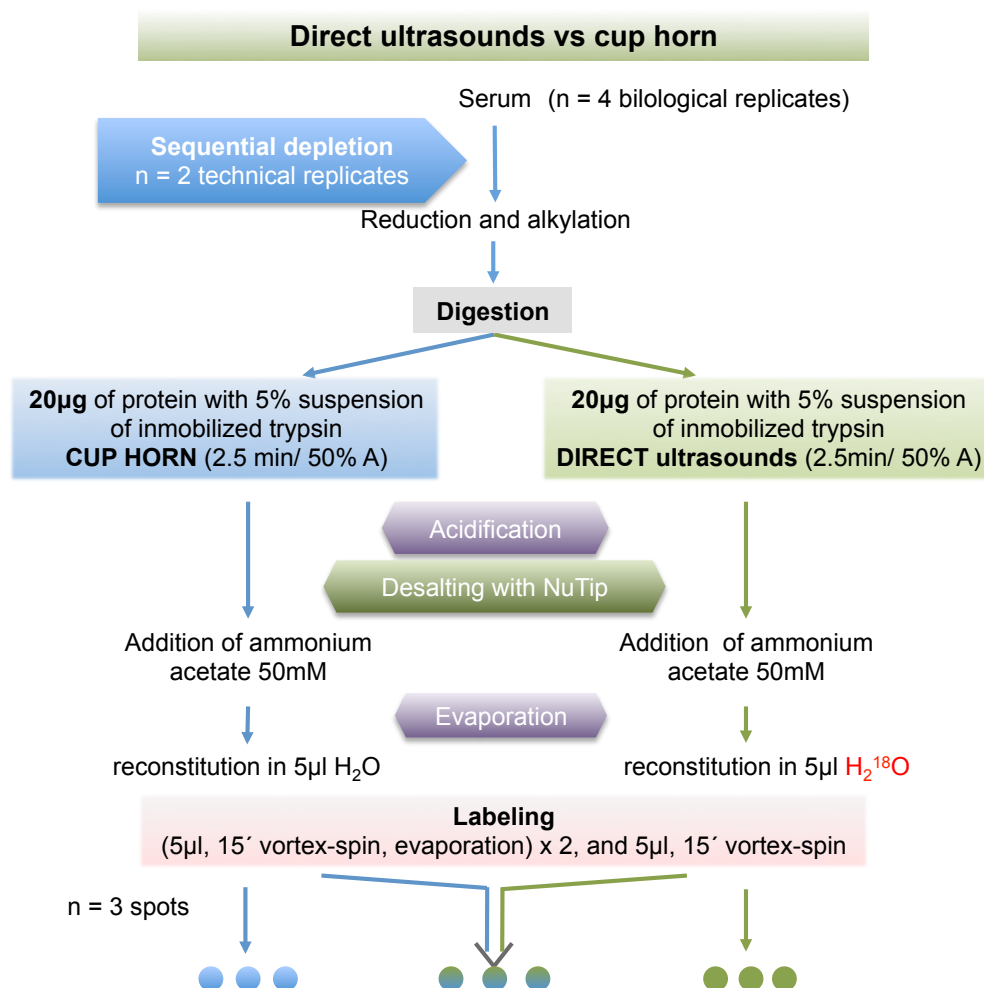


Figure 3SM. Efficiency of the double ^{18}O incorporation calculated in the peptides of the three different samples with four replicates.

Sample 1 (OA)		Sample 1 (OA)		Sample 3 (N)		Sample 2 (OA)	
m/z ($\Delta\text{Da} = 0$)	Efficiency \pm SD	m/z ($\Delta\text{Da} = 0$)	Efficiency \pm SD	m/z ($\Delta\text{Da} = 0$)	Efficiency \pm SD	m/z ($\Delta\text{Da} = 0$)	Efficiency \pm SD
972,55	100 \pm 0	1753,03	95 \pm 11	1012,65	100 \pm 0	1012,65	100 \pm 0
1012,63	100 \pm 0	1779,83	72 \pm 5	1031,63	100 \pm 0	1031,64	100 \pm 0
1031,57	100 \pm 0	1797,97	100 \pm 0	1104,67	100 \pm 0	1160,71	100 \pm 0
1101,63	100 \pm 0	1805,99	49 \pm 16	1160,71	100 \pm 0	1201,69	100 \pm 0
1104,63	96 \pm 9	1833,97	15 \pm 12	1194,72	100 \pm 0	1226,67	32 \pm 3
1112,57	100 \pm 0	1857,99	98 \pm 5	1196,71	100 \pm 0	1283,71	96 \pm 1
1126,62	100 \pm 0	1871,07	100 \pm 0	1201,70	80 \pm 29	1301,78	95 \pm 1
1160,65	100 \pm 0	1876,01	100 \pm 0	1226,66	89 \pm 11	1314,80	100 \pm 0
1201,62	100 \pm 0	1883,97	97 \pm 2	1283,70	96 \pm 1	1352,82	100 \pm 0
1226,60	77 \pm 4	1905,96	100 \pm 0	1301,77	94 \pm 2	1400,80	100 \pm 0
1230,74	100 \pm 0	1906,93	99 \pm 2	1311,83	100 \pm 0	1411,81	100 \pm 0
1233,64	100 \pm 0	1918,04	100 \pm 0	1352,82	100 \pm 0	1488,85	100 \pm 0
1235,70	100 \pm 0	1928,03	94 \pm 7	1400,79	100 \pm 0	1645,98	100 \pm 0
1252,67	100 \pm 0	1946,09	62 \pm 8	1411,81	100 \pm 0	1692,03	100 \pm 0
1276,70	98 \pm 5	1994,06	100 \pm 0	1488,85	100 \pm 0	1707,99	100 \pm 0
1281,65	100 \pm 0	2016,14	96 \pm 2	1532,93	100 \pm 0	1716,90	98 \pm 2
1283,64	100 \pm 0	2037,16	100 \pm 0	1645,99	90 \pm 2	1724,12	100 \pm 0
1301,71	94 \pm 2	2081,08	100 \pm 0	1691,99	100 \pm 0	1742,97	97 \pm 1
1311,76	100 \pm 0	2084,12	100 \pm 0	1708,02	98 \pm 4	1753,12	100 \pm 0
1314,72	100 \pm 0	2113,05	99 \pm 2	1716,97	97 \pm 3	1798,07	100 \pm 0
1350,72	100 \pm 0	2178,04	100 \pm 0	1724,11	100 \pm 0	1876,14	100 \pm 0
1352,73	100 \pm 0	2185,11	92 \pm 11	1742,97	97 \pm 1	1883,13	100 \pm 0
1386,78	100 \pm 0	2186,11	69 \pm 6	1753,13	100 \pm 0	1884,09	97 \pm 1
1387,75	100 \pm 0	2203,15	100 \pm 0	1798,07	100 \pm 0	1907,03	100 \pm 0
1394,68	100 \pm 0	2259,21	100 \pm 0	1858,09	100 \pm 0	1946,20	53 \pm 9
1400,72	100 \pm 0	2350,17	87 \pm 9	1876,13	96 \pm 8	2016,25	95 \pm 1
1411,74	100 \pm 0	2381,27	100 \pm 0	1883,12	100 \pm 0	2059,13	100 \pm 0
1451,75	100 \pm 0	2451,28	100 \pm 1	1884,09	96 \pm 2	2081,21	100 \pm 0
1456,78	100 \pm 0	2555,35	100 \pm 0	1906,99	100 \pm 0	2113,16	100 \pm 0
1465,73	99 \pm 3	3140,58	100 \pm 0	1928,12	100 \pm 0	2163,24	92 \pm 13
1522,78	100 \pm 0			1933,10	100 \pm 0	2202,32	100 \pm 0
1529,80	98 \pm 2	Sample 3 (N)		1946,18	100 \pm 0	2273,35	100 \pm 0
1612,85	100 \pm 0	m/z ($\Delta\text{Da} = 0$)	Efficiency \pm SD	1994,17	100 \pm 0	2350,29	82 \pm 2
1645,90	100 \pm 0	2350,26	91 \pm 2	2016,24	94 \pm 2	2381,40	100 \pm 0
1657,83	-4 \pm 11	2372,27	97 \pm 5	2038,23	100 \pm 0	2403,41	100 \pm 0
1660,83	57 \pm 29	2381,38	80 \pm 29	2059,18	100 \pm 0	2451,41	100 \pm 0
1683,88	100 \pm 0	2403,40	100 \pm 0	2081,21	100 \pm 0	2555,49	100 \pm 0
1706,92	100 \pm 0	2451,39	98 \pm 1	2084,22	100 \pm 0	2966,60	93 \pm 3
1707,90	99 \pm 2	2555,45	100 \pm 0	2113,15	100 \pm 0	3349,80	100 \pm 0
1711,85	100 \pm 0	2618,44	90 \pm 9	2202,30	96 \pm 3	3350,77	100 \pm 0
1716,92	95 \pm 2	2966,59	100 \pm 0	2273,30	100 \pm 0	3433,83	100 \pm 0
1742,88	93 \pm 1	3140,69	100 \pm 0	2296,33	100 \pm 0		
		3349,71	100 \pm 0				

7.4. Anexo IV: Material suplementario del capítulo IV

Tabla de las proteínas identificadas mediante secuenciación en el espectrómetro de masas (MALDI-TOF/TOF) empleando el motor de búsqueda MASCOT.

Número de Acceso	Nombre	Abrev.	MW	Péptidos macheados	Secuencia cubierta %	Secuencia	Score
P02763	Alpha-1-acid glycoprotein 1	A1AG1	23725	5	21	K.WFYIASAFR.N R.YVGGQEHFAHLLILR.D R.YVGGQEHFAHLLILRDTK.T K.EQLGEFYEALDCLRI K.EQLGEFYEALDCLRIK.P.S	170
P19652	Alpha-1-acid glycoprotein 2	A1AG2	23873	4	20	K.WFYIASAFR.N R.YEGGREHVAHLLFLR.D R.EHVAHLLFLR.D K.EQLGEFYEALDCLCIPR.S	136
P01009	Alpha-1-antitrypsin	A1AT	46878	4	6	K.ITPNLAEEFAFSLYR.Q K.ITPNLAEEFAFSLYR.Q K.FNKPFFVFLMIEQNTK.S K.FNKPFFVFLMIEQNTK.S	225
P02765	Alpha-2-HS-glycoprotein	FETUA	40098	2	11	R.HTFMGVVSLGSPSGEVSHPR.K R.TVVQPSVGGAAAGPVVPPCPGR.I	210
P01023	Alpha-2-macroglobulin	A2MG	164614	3	2	K.AIGYLNTGYQR.Q K.HYDGSYSTFGER.Y K.DNSVHWERPQPK.A	138
P01019	Angiotensinogen	ANGT	53406	3	7	K.VLSALQAVQGLLVAAQGR.A K.QPFVQGLALYTPVVLPR.S K.QPFVQGLALYTPVVLPR.S	57
P02647	Apolipoprotein A-I	APOA1	30759	5	19	K.VQPYLDDFQKK.W K.LSPLGEEMR.D R.THLAPYSDELRL.Q R.LEALKENGGAR.L K.AKPALEDRL.Q	227
P02652	Apolipoprotein A-II	APOA2	11282	2	21	K.VKSPQLQAEAK.S K.SKEQLTPLIK.K	103
P06727	Apolipoprotein A-IV	APOA4	45371	3	7	R.LEPYADQLR.T R.QLTPYAQR.M R.RVEPYGENFNK.A	75
P02656	Apolipoprotein C-III	APOC3	10843	2	19	K.TAKDALSSVQESQVAQQAR.G K.DALSSVQESQVAQQAR.G	196
P05090	Apolipoprotein D	APOD	21547	2	17	K.CPNPPVQENFDVNKYLGR.W R.WYEIKIPTTFENGRC	148
P02749	Beta-2-glycoprotein 1	APOH	39584	1	4	R.VCPFAGILENGAVR.Y	48
A2IDD5	Coiled-coil domain-containing protein 78	CCD78	48777	1	1	R.VLEELRL.G	38
P00738	Haptoglobin	HPT	45861	4	19	K.AVGDKLPECEADDGCKPPEIAHGVEHSVR.Y R.VMPCPLPSKDYAEVGR.V K.YVMLPVADQDQCIR.H K.SPVGVQPILNEHTFCAGMSK.Y	313
P69905	Hemoglobin subunit alpha	HBA	15305	1	11	K.TYEPHFDSLHGSAQVK.G	84
P68871	Hemoglobin subunit beta	HBB	16102	3	32	K.SAVTALWGKVVNDEVGGEALGR.L R.LLVVYPWTQR.F R.LLGNVLVLCVLAHHEGK.E	108
P01876	Ig alpha-1 chain C region	IGHA1	38486	2	10	R.QEPSQGTTFFAVTSILR.V K.KGDTFSCMVGHEALPLAFTQK.T	78
P01857	Ig gamma-1 chain C region	IGHG1	36596	4	14	K.KCDKTHTCPPCAPELLGGPSVFLFPPKPK.D K.THTCPPCAPELLGGPSVFLFPPKPK.D R.VVSVLTVLHQDWLNGK.E R.VVSVLTVLHQDWLNGKEYK.C	313
P01859	Ig gamma-2 chain C region	IGHG2	36505	2	14	K.CCVECPCPPAPPVAGPSVFLFPPKPK.D R.VVSVLTVVHQDWLNGKEYK.C	84
P01860	Ig gamma-3 chain C region	IGHG3	42287	2	5	K.GPSVFLAPCSR.S K.SCDTPPPCPR.C	73
P01766	Ig heavy chain V-III region BRO	HV305	13332	1	15	-EVOLVESGGGLVQPGGSLR.L	119
P01765	Ig heavy chain V-III region TIL	HV304	12462	1	16	-EVOLVESGGGLVQPGGSLR.L	35
P01834	Ig kappa chain C region	IGKC	11773	2	32	K.TVAAPSVFIFPPSDEQLK.S K.SGTASVVCCLNNFYPR.E	139
P01619	Ig kappa chain V-III region B6	KV301	11742	1	16	-QIVLTESPTGLSLSPGER.A	44
P01620	Ig kappa chain V-III region SIE	KV302	11882	1	16	-EIVLTQSPGTLSLSPGER.A	65
P0CG04	Ig lambda-1 chain C regions	LAC 1	11512	1	14	K.YAASSYLSLTPEQWK.S	71
P01871	Ig mu chain C region	IGHM	49960	1	3	K.YVTSAPMPPEQAPGR.Y	72
Q8N239	Kelch-like protein 34	KLH34	71309	1	1	R.DRWTAAGALPRA	32
P01042	Kininogen-I	KNG1	72996	1	1	K.CPGRPWK.S	35
P02750	Leucine-rich alpha-2-glycoprotein	A2GL	38382	2	6	R.YLFLNGNKLAR.V R.VAAGAFQGLR.Q	57
Q6B016	Lysine-specific demethylase 4D	KDM4D	59193	1	2	R.AQELTLQTPAK.R	35
Q6ZNJ1	Neurobeachin-like protein 2	NBEL2	305405	2	0	R.NAVAAFLMLLR.N + Oxidation (M) R.NAVAAFLMLLR.N	27
P20742	Pregnancy zone protein	PZP	165242	1	1	K.AGAFCLSEDAGLGISSTASLR.A	122
Q8NA82	Probable E3 ubiquitin-protein ligase MARCH10	MARHA	91310	1	1	K.VDSEYQACLR.R	28

P00734	Prothrombin	THRB	71475	1	2	R.TATSEYQTFNPR.T	84
Q86YV0	RAS protein activator like-3	RASL3	112741	2	2	K.HTPLISKSQLR.S R.ALTEQQEQLR.G	34
Q8NHV9	Rhox homeobox family member 1	RHXF1	20700	1	4	R.ELMLANELR.A + Oxidation (M)	43
P02787	Serotransferrin		79280	1	2	R.SAGWNIPIGLLYCDLPEPR.K	41
P02768	Serum albumin	ALBU	71317	4	12	K.SLHTLFGDKLCTVATLR.E R.LVRPEVDVMCTAFHDNEETFLKK.Y K.VFDEFKPLVEEPQNLK.Q K.EFNAETFTFHADICTLSEKER.Q	95
P02735	Serum amyloid A-1 protein	SAA1	13581	1	11	R.EANYIGSDKYFHAR.G	48
P35542	Serum amyloid A-4 protein	SAA4	14854	1	7	K.SNEKAEWGR.S	54
Q13033	Striatin-3	STNR3	87554	1	1	K.YTLRSHFDGVR.A	32
Q9Y4K3	TNF receptor-associated factor 6	TRAF6	61244	1	2	R.EGFQPRSTDAGV.-	44
Q92994	Transcription factor IIIB 90 kDa subunit	TF3B	74763	1	1	R.EPIQASTAR.E	33
P02766	Transferrin	TTHY	15991	3	25	R.GSPAINVAVHVFR.K R.GSPAINVAVHVFR.K K.ALGISPFHEHAEVVFTANDSGPR.R	91
Q147U7	Single-pass membrane and coiled-coil domain-containing protein 1 (Uncharacterized protein C3orf43)	SMCO1 (CC043)	24753	1	4	K.NQALQDSLRLA	34
Q9BQ65	U6 snRNA phosphodiesterase (UPF0406 protein C16orf57)	USB1 (CP057)	30420	1	5	R.MKVFHLSLSQSVVLR.H	47
Q8N6Y0	Usher syndrome type-1C protein-binding protein 1	USBP1	76364	1	1	R.ELEQVAQKGR.A	31

7.5. Anexo V: Artículo V. Identification of a Panel of Novel Serum Osteoarthritis Biomarkers

Journal of
proteome
research

ARTICLE

pubs.acs.org/jpr

Identification of a Panel of Novel Serum Osteoarthritis Biomarkers

Patricia Fernández-Puente,[†] Jesús Mateos,[†] Carolina Fernández-Costa,[†] Natividad Oreiro,[†] Carlos Fernández-López,[†] Cristina Ruiz-Romero,^{*,†} and Francisco J. Blanco^{*,†,‡}

[†]Osteoarticular and Aging Research Lab, Proteomics Unit-ProteoRed/ISCIII, INIBIC-Complejo Hospitalario Universitario A Coruña, 15006-A Coruña, Spain

[‡]CIBER-BBN Instituto Salud Carlos III, A Coruña, Spain

[§]Supporting Information

ABSTRACT: Osteoarthritis (OA) is the most common rheumatic pathology. Because currently available diagnostic methods are limited and lack sensitivity, the identification of new specific biological markers for OA has become a focus. The purpose of this study was to identify novel protein biomarkers for moderate and severe OA in serum. Sera were obtained from 50 moderate OA patients, 50 severe OA patients, and 50 nonsymptomatic controls. Serum protein levels were analyzed using isobaric tags for relative and absolute quantitation (iTRAQ) and matrix-assisted laser desorption/ionization (MALDI)-TOF/TOF mass spectrometry. We identified 349 different proteins in the sera, 262 of which could be quantified by calculation of their iTRAQ ratios. Three sets of proteins were significantly ($p < 0.05$) changed in OA samples compared to controls. Of these, 6 were modulated only in moderate OA, 13 only in severe OA and 7 in both degrees. Although some of these proteins, such as cartilage oligomeric matrix protein, have a previously reported putative biomarker value for OA, most are novel biomarker candidates for the disease. These include some complement components, lipoproteins, von Willebrand factor, tetranectin, and lumican. The specificity and selectivity of these candidates need to be validated before new molecular diagnostic or prognostic tests for OA can be developed.

KEYWORDS: osteoarthritis, biomarker, proteomics, iTRAQ, serum

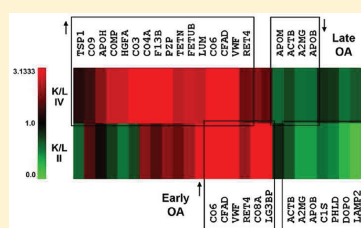


Table 1. Characteristics of the Osteoarthritis (OA) Patients and Controls in This Study

	control	moderate OA (K/L II)	severe OA (K/L IV)
iTRAQ-labeled samples	114	115	116
Number of patients	50	50	50
Age ^a	70.2 ± 7.70	72.4 ± 6.13	74.4 ± 5.32
% Female	42	64	84
Protein ^b	56.11 ± 5.64	55.5 ± 4.55	56.22 ± 6.74
Pool 1			
Number of patients	10	10	10
Age	69.8 ± 8.32	72.2 ± 7.25	68.33 ± 2.08
Protein	55.93 ± 1.30	55.1 ± 6.08	57.61 ± 9.52
Pool 2			
Number of patients	10	10	10
Age	72.7 ± 7.59	72.7 ± 5.79	74.4 ± 6.87
Protein	54.59 ± 9.38	56.8 ± 4.8	54.31 ± 4.43
Pool 3			
Number of patients	10	10	10
Age	71.5 ± 9.19	72.7 ± 6.29	76 ± 7.76
Protein	59.44 ± 0.16	57.4 ± 0.59	55.59 ± 7.78
Pool 4			
Number of patients	10	10	10
Age	70.10 ± 7.19	72.4 ± 5.17	75.1 ± 7.75
Protein	57.12 ± 0.62	53.6 ± 1.14	58.23 ± 6.30
Pool 5			
Number of patients	10	10	10
Age	66.9 ± 6.19	72 ± 7.18	74.4 ± 5.32
Protein	53.38 ± 5.64	54.4 ± 1.14	55.28 ± 4.85

^a Age shown as years with mean ± SD. ^b Protein shown in $\mu\text{g}/\mu\text{L}$ with mean ± SD.

according to the American College of Rheumatology (ACR) criteria, and knee and hip radiographs from all the participants were classified from grade I to grade IV according to the Kellgren and Lawrence (K/L) scoring system.¹³ The patients were of both genders and ages ranged from 58 to 90 years. A population of 50 samples each from K/L grade II (mild/early OA), K/L grade IV (late/advanced OA), and nonsymptomatic controls (K/L grade 0) was utilized for this study. Prior to proteomic analysis, the serum samples were grouped into five pools of 10 samples each to reduce individual and biological variability. Table 1 summarizes the characteristics of each pool.

Depletion of the 20 Most Abundant Serum Proteins

Prior to immunodepletion, crude serum samples (150 μL) were filtered through 0.22 μm cellulose acetate membrane tube filters (CLS8160, Costar Spin-X centrifuge, Corning Life Sciences, Amsterdam, The Netherlands) at 14 000 $\times g$ for 45 min. The filtered serum (90 μL) was then injected into a high performance liquid chromatography system (1200 Series, Agilent Technologies, Palo Alto, CA), equipped with a ProteoPrep20 Plasma Immunodepletion LC Column (Sigma-Aldrich, St. Louis, MO). Buffer A (ProteoPrep equilibration buffer, (PBS), Sigma) was used in the equilibration and washing steps. A ProteoPrep elution buffer (0.1 M Glycine-HCl, pH 2.5 with Octyl β -D-glucopyranoside, Sigma) was used for elution of the bound and highly abundant proteins from the column. The chromatograms of the pools were monitored using a UV detector at 280 nm at room temperature according to the protocol

supplied by the manufacturer (Sigma-Aldrich). The flow-through fractions containing low-abundance proteins were collected and concentrated using an Amicon ultra centrifuge (3 kDa M_c cutoff, Millipore Corporation, Billerica, MA) at 4000 $\times g$ for 1 h. The protein concentration of whole and depleted serum before and after being injected onto the immunodepletion column was determined using a nanoDrop instrument (Thermo Fisher Scientific, Waltham, MA) and stored at -20°C .

Protein Digestion and Peptide iTRAQ Labeling

Equal amounts of depleted serum proteins (50 μg) from each pool were denatured with 2% sodium dodecyl sulfate (SDS) in 1 M triethylammonium bicarbonate (TEAB) (ABSciex, Foster City, CA). The samples were then reduced for 1 h at 60°C using 50 mM tris-(2-carboxyethyl) phosphine (TCEP) (ABSciex), and cysteine-blocked with 84 mM iodoacetamide (Sigma-Aldrich) at room temperature in the dark for 30 min. The proteins were digested with spectrometry grade trypsin (Gold Mass, Promega, Madison, WI) at a concentration of 1:20 trypsin/protein for 14 h at 37°C . Each peptide solution was labeled for 1.5 h at room temperature using the iTRAQ reagents previously reconstituted in 70 μL of ethanol, following the iTRAQ Reagents Multiplex kit-Plasma Protocol (ABSciex). The samples were labeled with iTRAQ reagents as follows: controls, 114; K/L grade II, 115; K/L grade IV, 116. The reaction was stopped by adding deionized water, and the labeled samples (114, 115 and 116, one from each group) were combined. The mixture was desalted using reversed

phase matrix porosR2 homemade columns (ABSciex) prior to LC–MS analysis.

Liquid Chromatography (LC)

The peptide mixture was separated by off-line reversed-phase chromatography at a basic pH (pH 10) to lower the complexity of the mixture. Dried labeled peptides were resuspended in 160–210 μ L of buffer A (10 mM ammonium hydroxide, 5% acetonitrile) and 150–200 μ L were injected into an HP 1200 system (Agilent). The separation was performed on a C18 reversed-phase column (Zorbax extend C18, 100 \times 2.1 mm id, 3.5 μ m, 300 Å; Agilent). The flow rate used was 0.2 mL/minute and the gradient was 100% buffer A for 10 min, 0–90% buffer B (10 mM ammonium hydroxide, 90% acetonitrile) for 60 min, then 100% buffer B for eight minutes, and finally, 100% A for 20 min. The chromatogram was produced using a UV Detector at 214 nm. Sixty fractions were collected every 90 s using a Gilson FC203B fraction collector (Gilson, Middleton, WI). Several fractions were pooled postcollection based on the peak intensity of the UV trace, yielding a total of 25 samples per LC run. Each fraction was dried in a vacuum concentrator and stored at -20°C for the next step of analysis.

The dried peptide fractions were dissolved in 0.1% trifluoroacetic acid (TFA) and 2% acetonitrile; 5 μ L of this sample were injected onto a capillary trap column (0.5 \times 2 mm, Michrom Bioresources, Auburn, CA) at a flow rate of 15 μ L/minute. Peptides were desalted for 10 min and loaded onto a C18 column (Integragit C18, Proteopec II, 75 μ m id, 10.2 cm, 5 μ m, 300 Å; New Objective, Woburn, MA) at a constant flow rate of 350 nL/minute to perform the separation. The peptides were separated using linearly increasing concentration of acetonitrile in buffer B. The reversed-phase fractions were collected and mixed with matrix α -cyano-4-hydroxycinnamic acid (Sigma-Aldrich) at 3 mg/mL in 0.1% TFA in 70% acetonitrile at a flow rate of 1.2 μ L/minute, using the Sun Collect MALDI Spotter/Micro Collector (SunChrom Wissenschaftliche Geräte GmbH, Germany). Fractions collected every 15 s were then spotted onto a matrix-assisted laser desorption/ionization (MALDI) plate for MS analysis.

MS/MS Analysis

Data acquisition was accomplished in a positive ion mode using a 4800 analyzer equipped with time-of-flight (TOF)/TOF ion optics (ABSciex) and 4000 Series Explorer software version 3.5.1 (ABSciex). MS spectra from m/z 800–4000 were acquired for each fraction using 1500 laser shots processed with internal calibration. We used an angiotensin 3 fmol/spot for the internal standard at a m/z of 1046.50 diluted in the matrix, using 3800 as the laser intensity. After screening of all LC-MALDI sample positions in MS positive reflector mode the fragmentation of automatically selected precursors was performed at a collision energy of 1 kV with CID gas (air). Up to 12 of the most intense ion signals per spot position with signal/noise (S/N) above 80 were selected as precursors for MS/MS acquisition, excluding common trypsin autolysis peaks and matrix ion signals. The number of shots was 1800 for MS/MS, and the laser intensity was set to 4700. A second MS/MS was acquired excluding the precursors selected in the previous MS/MS run. Precursors with a S/N above 25–40 were selected to identify proteins that were not identified in the previous run.

Data Analysis

Protein identification used ProteinPilot software v.3.0 (ABSciex). The Paragon algorithm in ProteinPilot software¹⁴ served as the

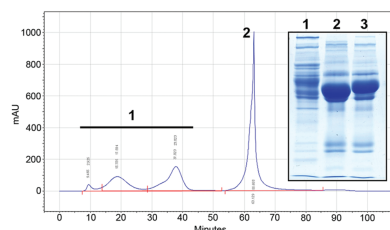


Figure 1. Depletion of 20 of the most abundant proteins in serum prior to proteomic analysis. A representative chromatogram of serum depletion is shown, together with an SDS-PAGE image of the different protein fractions. (1) Flow-through fraction (depleted serum); (2) bound fraction (high abundant proteins); (3) crude serum.

default search program with trypsin as the digestion agent and iodoacetamide as a fixed modification of cysteine. Biological modifications were programmed in the algorithm. Each MS/MS spectrum was searched in the Uniprot/Swissprot database (downloaded in May 2010) for *Homo sapiens*. Only proteins identified with at least 95% confidence, or a Prot Score (ProteinPilot Software protein confidence measure) of at least 1.3 were reported (Supplementary Table 1, Supporting Information). The ProteinPilot software also calculated a confidence percentage, the Unused score, which reflects the probability of a hit being a “false positive,” at a 95% confidence. While this software automatically accepts all peptides with an identification level >1%, only proteins having at least one peptide above the 95% confidence level were initially recorded. Searches against a concatenated database containing both forward and reversed sequences allowed the false discovery rate to be kept at 1%. Data were normalized for loading error by bias and the background correction was calculated using the Pro Group algorithm (ABSciex). The results obtained from ProteinPilot were exported to Microsoft Excel for further analyses. The PSPEP program was used independently to calculate false discovery rates.¹⁵

Statistical Analysis

The statistical significance of the differences between means was determined using the two-tailed *t*-test, accepting $p \leq 0.05$ to be significant. For proteomic analysis, the normalization tools and statistical package from ProteinPilot software were employed. A $p \leq 0.05$ or a ratio ≥ 1.3 or ≤ 0.7 was considered significant. Where appropriate, results are expressed as mean \pm standard error.

RESULTS AND DISCUSSION

This study has identified a panel of 28 different proteins that are altered in serum from OA patients when compared to controls using a shotgun quantitative proteomics approach. The strategy used is based on differential labeling of peptides with iTRAQ reagents prior to their separation and analysis by multidimensional LC coupled to MS. Recent analogous rheumatology studies have attempted to discover novel RA^{9,10} or SLE¹¹ biomarkers in serum or peripheral blood mononuclear cells (PMBCs), but this study is the first to focus this approach on OA. The population size employed in this shotgun proteomics study ($n = 150$, pooled in 5 separate experiments) is much larger

Table 2. Osteoarthritis (OA) Protein Biomarker Candidates Identified in This Work^a

accession number ^b	abbrev. ^b	protein name	moderate OA/Control ^c	PVal ^d	severe OA/control ^c	PVal ^d
Increased in OA						
P02749	APOH	Beta-2-glycoprotein 1	1.1376	0.126	2.1478	0.0049
P49747	COMP	Cartilage oligomeric matrix protein	0.7586	0.187	2.704	0.0475
P05160	F13B	Coagulation factor XIII B chain	1.7219	0.4415	3.5645	0.0285
P01024	CO3	Complement C3	0.7112	0.1977	8.091	0.0148
P0COL4	CO4A	Complement C4-A	2.1281	0.1681	4.7863	0.0319
P13671	CO6	Complement component C6	4.2073	0.0002	4.7863	<0.0001
P07357	CO8A	Complement component C8 alpha chain	3.1333	0.0062	2.1086	0.0704
P02748	CO9	Complement component C9	1.7061	0.0752	1.5417	0.0277
P00746	CFAD	Complement factor D	3.767	0.0457	3.8726	0.0114
Q9UGM5	FETUB	Fetuin-B	2.355	0.1888	2.704	0.0256
Q08380	LG3BP	Galectin-3-binding protein	4.8306	0.0042	1.8535	0.1604
Q04756	HGEFA	Hepatocyte growth factor activator	0.5445	0.2971	2.729	0.0519
P51884	LUM	Lumican	2.9923	0.1075	2.6303	0.0067
P20742	PZP	Pregnancy zone protein	2.1677	0.1011	3.7325	0.0004
P02753	RET4	Retinol-binding protein 4	2.2699	0.0492	1.9588	0.0236
P05452	TETN	Tetranectin	1.9409	0.0941	2.6062	0.0459
P04275	VWF	von Willebrand factor	5.5463	<0.0001	5.2966	<0.0001
P07996	TSP1	Thrombospondin-1	0.6081	0.5121	1.3428	0.0087
Decreased in OA						
P60709	ACTB	Actin, cytoplasmic 1	0.6855	0.0023	0.7178	0.0019
P01023	A2MG	Alpha-2-macroglobulin	0.3767	0.0072	0.5546	0.0105
P04114	APOB	Apolipoprotein B-100	0.3802	<0.0001	0.5702	<0.0001
O95445	APOM	Apolipoprotein M	0.8472	0.0883	0.5861	0.0228
P09871	C1S	Complement C1s subcomponent	0.5546	0.0027	0.8166	0.1038
P09172	DOPO	Dopamine beta-hydroxylase	0.2582	0.0429	0.5495	0.1562
P13473	LAMP2	Lysosome-associated membrane glycoprotein 2	0.078	0.0437	0.6486	0.089
P80108	PHLD	Phosphatidylinositol-glycan-specific phospholipase D	0.4699	0.0478	0.787	0.8835

^a Numbers in bold depict statistically significant ($p < 0.05$) changes. The number of unique peptides detected by ProteinPilot, total score and sequence coverage for each protein identified are detailed in Supplementary Table 1, Supporting Information. ^b Swiss-Prot accession number and abbreviation. ^c Average ratio, quantified by ProteinPilot 3.0 software. The iTRAQ ratios are quantified relative to the protein abundance in OA patients compared to controls. ^d Determined by ProteinPilot 3.0 software.

than used in previous works, providing increased consistency of results.

Sample Preparation and Processing for Quantitative Proteomics Analyses

Very recently, De Seny and colleagues¹⁶ have reported a proteomics-based discovery of four novel OA biomarkers using a Surface Enhanced Laser Desorption/Ionization (SELDI)-based technique.¹⁷ This strategy, although providing a convenient throughput that enables screening of a large number of samples, has the limitation of reducing the analysis to a specific proteome subset, only those proteins that remain bound to the selected protein chip according to their biochemical properties; this leads to "MS peaks" whose corresponding protein cannot always be identified. Unlike the SELDI technique, the iTRAQ-based method theoretically covers the entire proteomic profile of the samples. This results in a challenging issue for biomarker identification and quantification, due to the high complexity and extreme dynamic range (10 orders of magnitude) of serum proteins.¹⁸ Therefore, we utilized a targeted protein depletion step employing affinity chromatography to remove the 20 most abundant plasma proteins that mask those less abundant. This strategy, although might lead to the concomitant removal of

some nontargeted proteins that may be of potential interest, has proven to be essential for enhancing detection sensitivity and improving protein profiling capacities of human serum.¹⁹ From the different depletion methods tested, column immunodepletion showed the lowest removal of nontargeted proteins and proved to be the most promising depletion approach and a reliable method for serum preparation prior to proteomic studies.²⁰

Figure 1 shows a representative chromatogram of the depletion process. Two cycles of depletion were performed for each serum, yielding highly reproducible chromatograms. Using this technique, we were able to obtain $226 \pm 85 \mu\text{g}$ of proteins from each depletion of $100 \mu\text{L}$ of crude serum, averaging $55.56 \pm 1.6 \mu\text{g}/\mu\text{L}$. Depleted and crude serum proteins were digested and analyzed using MALDI-MS (data not shown). Only two of the top 20 proteins—apolipoprotein A-1 and complement C3—were identified in the depleted sample, demonstrating the efficiency of the depletion. This technique enabled identification of such low concentration serum proteins as insulin-like growth factor binding protein 3, mannose-binding protein C and sex hormone-globulin, present at concentrations in the range of ng/mL .^{21,22} All of these were identified by more than one peptide

at 95% confidence. The reduction in the limit of detection obtained demonstrates the usefulness of our approach.

Protein Identification in Serum

A quantitative proteomic analysis was carried out on the depleted samples. Proteins in each pool (Table 1) were digested and differentially labeled with iTRAQ reagents. The labeled samples were then mixed and the peptides in the mixtures were resolved using two-dimensional LC prior to MS analysis for the identification and relative quantification of the corresponding proteins. Biological and technical replicate samples were used to increase the reliability of the iTRAQ technique for relative quantitation.²³ A total of 349 proteins could be identified in the sera after combining the results from all five pools. 204 of these proteins were identified with two or more peptides with 95% confidence, while another 145 were detected from a single peptide assignment (see Supplementary Table 1, Supporting Information). The David homology tool²⁴ was used to assign the molecular function, biological process and cellular compartment of the identified proteins (Supplementary Figure 1, Supporting Information). The identified proteins were involved in a wide range of biological processes, including immune response, proteolysis and inflammatory response, and are implicated in such different molecular functions as calcium carbohydrate binding, enzyme inhibition and peptidase activity. As expected, most are predicted to be extracellular.

We were able to identify some components of the cartilage ECM, such as cartilage oligomeric matrix protein (COMP), lumican (LUM), cartilage acidic protein 1 (CRAC1) or proteoglycan 4 (PRG4), and also proteins previously suggested as biomarkers for OA (fibronectin, thrombospondins, gelsolin) or rheumatoid arthritis (such as calgranulins S100 A8/A9). However, the relative number of proteins directly related with the disease that we found was very low. This is due to the fact that, despite the depletion of the top 20 most abundant proteins in serum, there are still a number of other typical serum proteins (i.e., complements, apolipoproteins and others) that still mask the identification of tissue-derived molecules.

Moreover, the iTRAQ technique (and other isobaric tagging methods for quantitative analyses) is a shotgun proteomics approach. In these strategies there is always a bias toward proteins that are abundant in the sample, since they are more likely to produce high MS peaks and are therefore easily selected for fragmentation and quantification. In addition, iTRAQ is known to display an underestimation tendency affected by interfering factors that depend on sample complexity, although provides accurate quantification across 2 orders of magnitude.²⁵ These idiosyncrasies inherent to the followed strategy would have undoubtedly conditioned the results.

Quantitative Proteomic Analysis

The relative abundance of 262 proteins was quantifiable by calculation of the corresponding iTRAQ label ratios (Supplementary Table 1, Supporting Information); 186 were quantified with more than 1 peptide at 95% of confidence and another 76 were quantified with statistical significance ($p < 0.05$). Previous studies have shown cutoff ratios indicating significant changes at the protein level to be 1.2 and 0.8^{26,27} when biological replicates are employed and three peptides are quantified. Therefore, for our study we considered those proteins exhibiting iTRAQ ratios greater than ($>$) 1.3 or less than ($<$) 0.7 to be differentially abundant when found in all five pools with a $p < 0.05$ and more than one quantified peptide with a

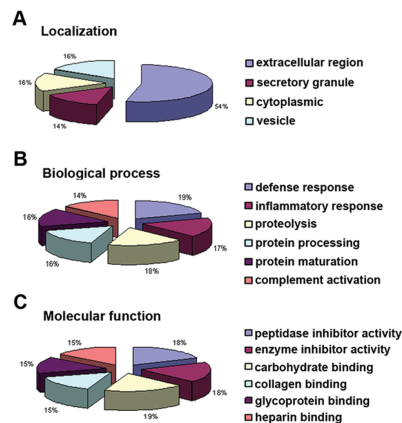


Figure 2. Data mining of the set of osteoarthritis (OA) serum biomarker candidates. (A) Predicted localization. (B) Predicted biological process. (C) Predicted molecular function.

confidence of 95%. Protein profiles between the OA patients (moderate or severe) and nonsymptomatic controls were compared using these criteria, leading to the identification of 26 proteins altered in OA: 18 of them were increased (OA/Control > 1.3 , $p < 0.05$) and 8 were decreased (OA/Control < 0.7 , $p < 0.05$). Table 2 summarizes the quantitative data for these proteins. Most are extracellular proteins involved in immune and inflammatory responses, proteolysis and protein processing (Figure 2).

The use of relative quantification enabled the characterization of a novel panel of moderate and/or severe OA biomarkers, illustrated by the hierarchical clustering seen in Figure 3. Proteins in this panel can be grouped into three subsets. In the first subset, only six proteins, two of which were increased and four decreased compared to controls, were specifically altered in moderate OA and are therefore interesting candidates for OA diagnosis or prognosis studies; these included the glycoproteins LAMP2 (lysosome-associated membrane glycoprotein 2), PHLN (phosphatidylinositol-3-OH kinase class I-related protein), and LG3BP (galectin-3-binding protein). Interestingly, the latter is a cell-adhesive protein of the extracellular matrix that binds integrins, collagens and fibronectin,²⁸ which has been previously proposed as marker in cancer and also in age-related macular degeneration.²⁹ In the second set, seven proteins were modulated in both moderate and severe degrees of OA. One of these was VWF (von Willebrand factor), a protein essential for homeostasis. Although a role for VWF as a marker of endothelial dysfunction has been previously defined for RA³⁰ and other chronic synovitis syndromes,³¹ this is the first time a significant increase in this protein is described in OA patients. Interestingly, we detected this increase already in K/L grade II samples, which should exhibit a lower inflammation component than grade IV. The discriminative value of this marker remains to be evaluated.

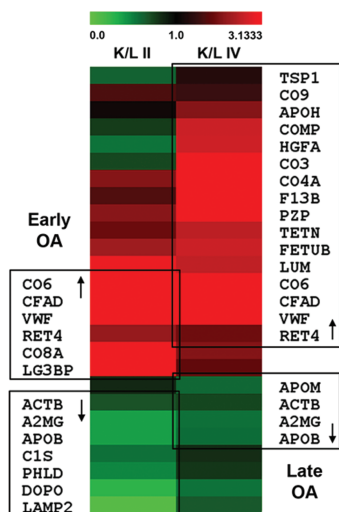


Figure 3. Hierarchical clustering representation of the altered serum protein profiles identified in osteoarthritis (OA) patients using a shotgun quantitative proteomics approach. The panel of moderate OA biomarker candidates is on the left and the panel of severe OA biomarkers is on the right. Protein abbreviations are found in Table 2.

Finally, the third set of proteins includes 13 where changes are seen only in severe OA. These would theoretically have less value as markers for early diagnosis. In this group (see Table 2), we detected increases of several immune response proteins, including complement components indicative of the response to extreme tissue degeneration characteristic of grade IV OA joints. Interestingly, the only two previously characterized cartilage-related proteins identified in this study as being altered in the OA sera, cartilage oligomeric matrix protein (COMP) and lumican (LUM), were found in this group. COMP is the only protein in our panel previously reported to be a biomarker for OA (reviewed in³²). COMP was somewhat, but statistically insignificantly, decreased in K/L II samples. LUM showed a nonsignificant ($p = 0.1$) tendency to increase in moderate degrees that might prove significant with a larger sample size. LUM is a member of the small leucine-rich proteoglycans (SLRPs) family, which plays important linking, shape-determining and matrix-organizing roles in musculoskeletal tissues.³³ A role has been recently proposed for this family in the regulation of complement activation in diseases involving the extracellular matrix, particularly those characterized by chronic inflammation, such as rheumatoid arthritis or osteoarthritis.³⁴ The increased fragmentation of SLRPs reported in degenerated human cartilage³⁵ could be the source for the increased amount of this protein found in OA patient's sera using our proteomics approach.

Altogether, our work has led to the characterization of a panel of proteins altered in the serum of OA patients, two of them being specific components of the cartilage ECM (one of them, LUM, not previously described as putative biomarker), which appear to be the most promising candidates *a priori*. This relatively low number cast doubt on the usefulness of shotgun proteomics approaches on plasma/serum samples, as hypothesis-free analyses will undoubtedly bear a "noise" from a number of ubiquitous proteins that can be altered during the progression of several unspecific processes. These might not be directly related with the pathology of interest, but with the general course of inflammation, immune response, cellular death/proliferation, etc. Nevertheless, performing the iTRAQ analysis directly on serum has the advantage of profiling the same fluid that will preferentially be the subject of verification and validation assays (due to the characteristics pointed out in the Introduction).

CONCLUSIONS

To summarize, we identified a panel of new OA serum protein biomarkers, some of which have an indicator value for moderate stages of the disease. Further analyses using a larger number of patients are needed to determine their sensitivity and specificity for OA and to evaluate the clinical benefit of using this panel for diagnosis or prognosis analyses.

ASSOCIATED CONTENT

Supporting Information

Supplementary Table 1 and Figure 1. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Unidad de Investigación del Envejecimiento Osteoarticular, INIBIC-Complejo Hospitalario, Universitario a Coruña C/Xubias, 84, 15006-A Coruña, Spain. Tel: 34-981-176399. Fax: 34-981-176398. E-mail: cristina.ruiz.romero@sergas.es or francisco.blanco.garcia@sergas.es.

ACKNOWLEDGMENT

We thank Ms. P. Cal Purniños for her expert secretarial assistance and the Rheumatology Service of CH Universitario A Coruña for providing serum samples. This study was supported by grants from Fondo Investigación Sanitaria (PI-08/2028), Secretaría I+D+I Xunta de Galicia (PGIDIT06PXIC91617SPN) and Instituto de Salud Carlos III (CIBER-BBN CB06/01/0040) Spain, with participation of funds from FEDER (European Community). P.F.-P., J.M., and C.R.-R. were supported by grants from Fondo Investigación Sanitaria-Spain.

REFERENCES

- (1) Pritzker, K. Pathology of osteoarthritis. In *Osteoarthritis*; Brandt, K., Doherty, M., Lohmander, L., Eds.; Oxford University Press: New York, 1998; pp 50–61.
- (2) Spector, T. D.; MacGregor, A. J. Risk factors for osteoarthritis: genetics. *Osteoarthritis Cartilage* **2004**, *12* (Suppl A), S39–44.
- (3) Valdes, A. M.; Van Oene, M.; Hart, D. J.; Surdulescu, G. L.; Loughlin, J.; Doherty, M.; Spector, T. D. Reproducible genetic associations between candidate genes and clinical knee osteoarthritis in men and women. *Arthritis Rheum.* **2006**, *54* (2), 533–9.

- (4) Roos, E. M. Joint injury causes knee osteoarthritis in young adults. *Curr. Opin. Rheumatol.* **2005**, *17* (2), 195–200.
- (5) Woolf, A.; Pfleger, B. Burden of major musculoskeletal conditions. *Bull. World Health Org.* **2003**, *81*, 10.
- (6) Peat, G.; Thomas, E.; Duncan, R.; Wood, L.; Wilkie, R.; Hill, J.; Hay, E. M.; Croft, P. Estimating the probability of radiographic osteoarthritis in the older patient with knee pain. *Arthritis Rheum.* **2007**, *57* (5), 794–802.
- (7) Ruiz-Romero, C.; Blanco, F. J. Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis Cartilage* **2010**, *18* (4), 500–9.
- (8) DeSouza, L.; Diehl, G.; Rodrigues, M. J.; Guo, J.; Romaschin, A. D.; Colgan, T. J.; Siu, K. W. Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cCAT with multidimensional liquid chromatography and tandem mass spectrometry. *J. Proteome Res.* **2005**, *4* (2), 377–86.
- (9) Dwivedi, R. C.; Dhindsa, N.; Krokhin, O. V.; Cortens, J.; Wilkins, J. A.; El-Gabalawy, H. S. The effects of infliximab therapy on the serum proteome of rheumatoid arthritis patients. *Arthritis Res. Ther.* **2009**, *11*, R32.
- (10) Serada, S.; Fujimoto, M.; Ogata, A.; Terabe, F.; Hirano, T.; Iijima, H.; Shinzaki, S.; Nishikawa, T.; Ohkawa, T.; Iwahori, K.; Ohguro, N.; Kishimoto, T.; Naka, T. iTRAQ-based proteomic identification of leucine-rich alpha-2 glycoprotein as a novel inflammatory biomarker in autoimmune diseases. *Ann. Rheum. Dis.* **2010**, *69*, 770–4.
- (11) Wang, L.; Dai, Y.; Qi, S.; Sun, B.; Wen, J.; Zhang, L.; Tu, Z. Comparative proteome analysis of peripheral blood mononuclear cells in systemic lupus erythematosus with iTRAQ quantitative proteomics. *Rheumatol. Int.* **2010**, DOI: 10.1007/s00296-010-1625-9.
- (12) Rego, I.; Fernandez-Moreno, M.; Fernandez-Lopez, C.; Gomer-Reino, J. J.; Gonzalez, A.; Arenas, J.; Blanco, F. J. Role of European mitochondrial DNA haplogroups in the prevalence of hip osteoarthritis in Galicia, Northern Spain. *Ann. Rheum. Dis.* **2010**, *69*, 210–3.
- (13) Kessler, S.; Guenther, K. P.; Puhl, W. Scoring prevalence and severity in gonarthrosis: the suitability of the Kellgren & Lawrence scale. *Clin. Rheumatol.* **1998**, *17* (3), 205–9.
- (14) Shilov, I. V.; Seymour, S. L.; Patel, A. A.; Loboda, A.; Tang, W. H.; Keating, S. P.; Hunter, C. L.; Nuwaysir, L. M.; Schaeffer, D. A. The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* **2007**, *6*, 1638–55.
- (15) Ruppen, I.; Grau, L.; Orenes-Piñero, E.; Ashman, K.; Gil, M.; Algaba, F.; Bellmunt, J.; Sánchez-Carbayo, M. Differential protein expression profiling by iTRAQ-two-dimensional LC-MS/MS of human bladder cancer EJ138 cells transfected with the metastasis suppressor KISS-1 gene. *Mol. Cell. Proteomics* **2010**, *9* (10), 2276–91.
- (16) de Seny, D.; Sharif, M.; Fillet, M.; Cobraiville, G.; Meuwis, M. A.; Maree, R.; Hauzeur, J. P.; Wehenkel, L.; Louis, E.; Merville, M. P.; Kirwan, J.; Ribbens, C.; Malaise, M. Discovery and biochemical characterisation of four novel biomarkers for osteoarthritis. *Ann. Rheum. Dis.* **2011**, *70*, 1144–52.
- (17) Seibert, V.; Wiesner, A.; Buschmann, T.; Meuer, J. Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI TOF-MS) and ProteinChip technology in proteomics research. *Pathol. Res. Pract.* **2004**, *200* (2), 83–94.
- (18) Anderson, N. L.; Polanski, M.; Pieper, R.; Gatlin, T.; Tirumalai, R. S.; Conrads, T. P.; Veenstra, T. D.; Adkins, J. N.; Pounds, J. G.; Fagan, R.; Lobley, A. The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol. Cell. Proteomics* **2004**, *3* (4), 311–26.
- (19) Echan, L. A.; Tang, H. Y.; Ali-Khan, N.; Lee, K.; Speicher, D. W. Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* **2005**, *5* (13), 3292–303.
- (20) Bellei, E.; Bergamini, S.; Monari, E.; Fantoni, L. L.; Cuoghi, A.; Ozben, T.; Tomasi, A. High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids* **2011**, *40* (1), 145–56.
- (21) Song, X.; Bandow, J.; Sherman, J.; Baker, J. D.; Brown, P. W.; McDowell, M. T.; Molloy, M. P. iTRAQ experimental design for plasma biomarker discovery. *J. Proteome Res.* **2008**, *7* (7), 2952–8.
- (22) Ernault, E.; Bourreau, A.; Gamelin, E.; Guette, C. A proteomic approach for plasma biomarker discovery with iTRAQ labelling and OFFGEL fractionation. *J. Biomed. Biotechnol.* **2010**, *2010*, 927917.
- (23) Gan, C. S.; Chong, P. K.; Pham, T. K.; Wright, P. C. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J. Proteome Res.* **2007**, *6* (2), 821–7.
- (24) Huang, d. W.; Sherman, B. T.; Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4* (1), 44–57.
- (25) Ow, S. Y.; Salim, M.; Noirel, J.; Evans, C.; Rehman, I.; Wright, P. C. iTRAQ underestimation in simple and complex mixtures: “the good, the bad and the ugly”. *J. Proteome Res.* **2009**, *8* (11), 5347–55.
- (26) Williamson, A. J.; Smith, D. L.; Blinco, D.; Unwin, R. D.; Pearson, S.; Wilson, C.; Miller, C.; Lancashire, L.; Lacaud, G.; Kouskoff, V.; Whetton, A. D. Quantitative proteomics analysis demonstrates post-transcriptional regulation of embryonic stem cell differentiation to hematopoiesis. *Mol. Cell. Proteomics* **2008**, *7*, 459–72.
- (27) Unwin, R. D.; Pierce, A.; Watson, R. B.; Sternberg, D. W.; Whetton, A. D. Quantitative proteomic analysis using isobaric protein tags enables rapid comparison of changes in transcript and protein levels in transformed cells. *Mol. Cell. Proteomics* **2005**, *4*, 924–35.
- (28) Sasaki, T.; Brakebusch, C.; Engel, J.; Timpl, R. Mac-2 binding protein is a cell-adhesive protein of the extracellular matrix which self-assembles into ring-like structures and binds beta1 integrins, collagens and fibronectin. *EMBO J.* **1998**, *17* (6), 1606–13.
- (29) An, E.; Lu, X.; Flippin, J.; Devaney, J. M.; Halligan, B.; Hoffman, E. P.; Strumikova, N.; Csaky, K.; Hathout, Y. Secreted proteome profiling in human RPE cell cultures derived from donors with age related macular degeneration and age matched healthy donors. *J. Proteome Res.* **2006**, *5* (10), 2599–610.
- (30) Daza, L.; Aguirre, M.; Jimenez, M.; Herrera, R.; Bollain, J. J. Common carotid intima-media thickness and von Willebrand factor serum levels in rheumatoid arthritis female patients without cardiovascular risk factors. *Clin. Rheumatol.* **2007**, *26* (4), 533–7.
- (31) Ogdie, A.; Li, J.; Dai, L.; Paessler, M. E.; Yu, X.; Diaz-Torne, C.; Akmatov, M.; Schumacher, H. R.; Pessler, F. Identification of broadly discriminatory tissue biomarkers of synovitis with binary and multi-category receiver operating characteristic analysis. *Biomarkers* **2010**, *15* (2), 183–90.
- (32) Morozzi, G.; Fabbri, M.; Bellisai, F.; Pucci, G.; Galeazzi, M. Cartilage oligomeric matrix protein level in rheumatic diseases: potential use as a marker for measuring articular cartilage damage and/or the therapeutic efficacy of treatments. *Ann. N.Y. Acad. Sci.* **2007**, *1108*, 398–407.
- (33) Scott, J. E. Elasticity in extracellular matrix ‘shape modules’ of tendon, cartilage, etc. A sliding proteoglycan-filament model. *J. Physiol.* **2003**, *553* (Pt 2), 335–43.
- (34) Sjöberg, A. P.; Manderson, G. A.; Mörgelin, M.; Day, A. J.; Heinegård, D.; Blom, A. M. Short leucine-rich glycoproteins of the extracellular matrix display diverse patterns of complement interaction and activation. *Mol. Immunol.* **2009**, *46* (5), 830–9.
- (35) Melrose, J.; Fuller, E. S.; Roughley, P. J.; Smith, M. M.; Kerr, B.; Hughes, C. E.; Caterson, B.; Little, C. B. Fragmentation of decorin, biglycan, lumican and keratan is elevated in degenerated human meniscus, knee and hip articular cartilages compared with age-matched macroscopically normal and control tissues. *Arthritis Res. Ther.* **2008**, *10* (4), R79.

